



VERIFICATION OF TRANSLATION

I, Mariko Ogura

of 6th Fl., Kantetsu Tsukuba-Science-City Bldg. 1-1-1, Oroshi-machi, Tsuchiura, Ibaraki,
JAPAN

declare as follows:

1. That I am well acquainted with both the English and Japanese languages, and
2. That the attached document is a true and correct translation made by me to the best of my knowledge and belief of:-

Japanese Patent Application No. Hei 10-297409
Entitled: "Novel Hemopoietin Receptor Proteins"
Filed on October 19, 1998

March 11, 2005
Date

Mariko Ogura
Mariko Ogura

BEST AVAILABLE COPY

【Name of Document】 APPLICATION FOR PATENT

【Identification Number】 C2-004

【Filing Date】 October 19, 1998

【Addressee】 Esq. Commissioner of the Patent Office

【International Patent Classification】 C12N 15/12

【Title of the Invention】 Novel Hemopoietin Receptor Proteins

【Number of Claims】 17

【Inventor】

【Address】 c/o CHUGAI RESEARCH INSTITUTE FOR MOLECULAR
MEDICINE, INC.

153-2, Nagai, Niihari-mura, Niihari-gun, IBARAKI

【Name】 Hitoshi Nomura

【Inventor】

【Address】 c/o CHUGAI RESEARCH INSTITUTE FOR MOLECULAR
MEDICINE, INC.

153-2, Nagai, Niihari-mura, Niihari-gun, IBARAKI

【Name】 Masatsugu Maeda

【Applicant】

【Identification Number】 596102791

【Name or Appellation】 CHUGAI RESEARCH INSTITUTE FOR MOLECULAR
MEDICINE, INC.

【Representative】 Yoshiyuki Ohsugi

【Attorney】

【Identification Number】 100102978

【Patent Attorney】

【Name or Appellation】 Hatsushi Shimizu

【Nominated Attorney】

【Identification Number】 100108774

【Patent Attorney】

【Name or Appellation】 Kazunori Hashimoto

【Priority】

【Application Number】 Hei10-214720

【Filing Date】 June 24, 1998

【Payment】

【Registration Number】 041092

【Amount】 21000 yen

【List of Attached Documents】

【Name of Document】 Specification 1

【Name of Document】 Drawings 1

【Name of Document】 Abstract 1

【General Power of Attorney Number】 9716403

[Document Name] Specification

[Title of the Invention] NOVEL HEMOPOIETIN RECEPTOR PROTEINS

[Claims]

[Claim 1] A protein comprising the amino acid sequence from
5 the 1st amino acid Met to the 361st amino acid Ser of SEQ ID NO: 1,
or a protein comprising a modified amino acid sequence of said amino
acid sequence in which one or more amino acids have been deleted,
added and/or substituted with another amino acid and being functionally
equivalent to the protein comprising the amino acid sequence from
10 the 1st amino acid Met to the 361st amino acid Ser of SEQ ID NO: 1.

[Claim 2] A protein comprising the amino acid sequence from
the 1st amino acid Met to the 144th amino acid Leu of SEQ ID NO: 3,
or a protein comprising a modified amino acid sequence of said amino
acid sequence in which one or more amino acids have been deleted,
15 added and/or substituted with another amino acid and being functionally
equivalent to the protein comprising the amino acid sequence from
the 1st amino acid Met to the 144th amino acid Leu of SEQ ID NO: 3.

[Claim 3] A protein comprising the amino acid sequence from
the 1st amino acid Met to the 237th amino acid Ser of SEQ ID NO: 5,
20 or a protein comprising a modified amino acid sequence of said amino
acid sequence in which one or more amino acids have been deleted,
added and/or substituted with another amino acid and being functionally
equivalent to the protein comprising the amino acid sequence from
the 1st amino acid Met to the 237th amino acid Ser of SEQ ID NO: 5.

[Claim 4] A protein comprising the amino acid sequence from
25 the 1st amino acid Met to the 538th amino acid Ser of SEQ ID NO: 7,
or a protein comprising a modified amino acid sequence of said amino
acid sequence in which one or more amino acids have been deleted,
added and/or substituted with another amino acid and being functionally
equivalent to the protein comprising the amino acid sequence from
30 the 1st amino acid Met to the 538th amino acid Ser of SEQ ID NO: 7.

[Claim 5] A protein encoded by a DNA hybridizing to a DNA
comprising the nucleotide sequence of SEQ ID NO: 2, said protein being
functionally equivalent to a protein comprising the amino acid sequence
35 from the 1st amino acid Met to the 361st amino acid Ser of SEQ ID NO:
1.

[Claim 6] A protein encoded by a DNA hybridizing to a DNA comprising the nucleotide sequence of SEQ ID NO: 4, said protein being functionally equivalent to a protein comprising the amino acid sequence from the 1st amino acid Met to the 144th amino acid Leu of SEQ ID NO: 3.

[Claim 7] A protein encoded by a DNA hybridizing to a DNA comprising the nucleotide sequence of SEQ ID NO: 6, which is functionally equivalent to a protein comprising the amino acid sequence from the 1st amino acid Met to the 237th amino acid Ser of SEQ ID NO: 5.

[Claim 8] A protein encoded by a DNA hybridizing to a DNA comprising the nucleotide sequence of SEQ ID NO: 8, said protein being functionally equivalent to a protein comprising the amino acid sequence from the 1st amino acid Met to the 538th amino acid Ser of SEQ ID NO: 7.

[Claim 9] A fusion protein comprising the protein of any one of claims 1 to 8 and another peptide or polypeptide.

[Claim 10] A DNA encoding the protein of any one of claims 1 to 9.

[Claim 11] A vector comprising the DNA of claim 10.

[Claim 12] A transformant harboring the DNA of claim 10 in an expressible manner.

[Claim 13] A method of producing the protein of any one of claims 1 to 8, comprising the step of culturing the transformant of claim 12.

[Claim 14] A method of screening a compound that binds to the protein of any one of claims 1 to 9 comprising the steps of:

(a) contacting a test sample with the protein of any one of claims 1 to 9; and

(b) selecting a compound that comprises an activity to bind to the protein of any one of claims 1 to 9.

[Claim 15] An antibody that specifically binds to the protein of any one of claims 1 to 8.

[Claim 16] A method of detecting or measuring the protein of any one of claims 1 to 9 comprising the steps of contacting a test sample presumed to contain said protein with the antibody of claim 15, and detecting or measuring the formation of the immune complex

between the antibody and the protein.

[Claim 17] A DNA specifically hybridizing to a DNA comprising a nucleotide sequence of any one of SEQ ID NOs: 2, 4, 6, and 8 comprising at least 15 nucleotides, and comprising at least 15 nucleotides.

5 [Detailed Description of the Invention]

[0001]

[Technical Field of Industrial Application]

The present invention relates to novel hemopoietin receptor proteins, the encoding genes, and methods of production and uses
10 thereof.

[0002]

[Prior Art]

A large number of cytokines are known as humoral factors that are involved in the proliferation/differentiation of various cells, or
15 activation of differentiated mature cells, and also cell death. These cytokines have their own specific receptors, which are categorized into several families based on their structural similarities (Hilton D.J., in "Guidebook to Cytokines and Their Receptors" edited by Nicola N.A. (A Sambrook & Tooze Publication at Oxford University Press),
20 1994, p8-16).

[0003]

Compared to similarities between receptors, primary-structure homology is quite low between cytokines, and a significant amino acid homology cannot be seen even among cytokine members that belong to
25 the same receptor family. This explains the functional specificity of each cytokine, as well as similarities of cellular reactions induced by each cytokine.

[0004]

Representative examples of the above-mentioned receptor families
30 are the tyrosine kinase receptor family, hemopoietin receptor family, tumor necrosis factor (TNF) receptor family, and transforming growth factor β (TGF β) receptor family. Different signal transduction pathways have been reported to be involved in each of these families. Among these receptor families, many receptors of especially the
35 hemopoietin receptor family are expressed in blood cells and immunocytes, and their ligands, cytokines, are often termed as

hemopoietic factors or interleukins. Some of these hemopoietic factors or interleukins exist within blood and are thought to be involved in a systemic humoral regulation of hemopoietic or immune functions.

[0005]

5 This contrasts with the belief that cytokines belonging to other families are often involved in only topical regulations. Some of these hemopoietins can be taken as hormone-like factors, and conversely, representative peptide hormones such as the growth hormone, prolactin, or leptin receptors also belong to the hemopoietin receptor family.
10 Because of these hormone-like systemic regulatory features, it is anticipated that hemopoietin administration would be applied in the treatment of various diseases.

[0006]

Among the large number of cytokines, those that are actually being
15 clinically applied are, erythropoietin, G-CSF, GM-CSF, and IL-2. Combined with IL-11, LIF, and IL-12 that are being considered for clinical trials, and the above-mentioned peptide hormones such as growth hormone and prolactin, it can be envisaged that by searching among the above-mentioned various receptor families for a novel
20 cytokine that binds to hemopoietin receptors, it is possible to find a cytokine that can be clinically applied with a higher efficiency.

[0007]

As mentioned above, cytokine receptors have structural similarities between the family members. Using these similarities,
25 many investigations are being carried out aiming at finding novel receptors. Regarding the tyrosine kinase receptor especially, many receptors have already been cloned using its highly conserved sequence at the catalytic site (Matthews W. et al., Cell, 1991, 65 (7) p1143-52). Compared to this, hemopoietin receptors do not have a tyrosine
30 kinase-like enzyme activity domain in their cytoplasmic regions, and their signal transductions are known to be mediated through associations with other tyrosine kinase proteins existing freely in the cytoplasm.

[0008]

35 Though the binding site on receptors associating with these cytoplasmic tyrosine kinases (JAK kinases) is conserved between family

members, the homology is not very high (Murakami M. et al., Proc. Natl. Acad. Sci. USA, 1991, 88, 11349-11353). On one hand, the sequence that characterizes these hemopoietin receptors most well exists in the extracellular region, and especially the five amino acid
5 Trp-Ser-Xaa-Trp-Ser (where Xaa is an arbitrary amino acid) motif is conserved in almost all of the hemopoietin receptors. Therefore, novel receptors are expected to be obtained by searching novel family members using this sequence. In fact, this approach has already identified the IL-11 receptor (Robb, L. et al., J. Biol. Chem., 1996, 271 (23)
10 13754-13761), leptin receptor (Gainsford T. et al., Proc. Natl. Acad. Sci. USA, 1996, 93 (25) p14564-8) and the IL-13 receptor (Hilton D.J. et al., Proc. Natl. Acad. Sci. USA, 1996, 93 (1) p497-501).

[0009]

[Problems to Be Solved by the Invention]

15 The present invention provides a novel hemopoietin receptor protein, and the encoding DNA. The present invention also provides, a vector into which the DNA has been inserted, a transformant harboring the DNA, and a method of producing a recombinant protein using the transformant. It also provides a method of screening a compound that
20 binds to the protein.

[0010]

[Means to Solve the Problems]

Until now, the inventors have been trying to search for a novel receptor using an oligonucleotide encoding the Trp-Ser-Xaa-Trp-Ser
25 motif as a probe by plaque hybridization, RT-PCR method, and so on. However, because of reasons such as the oligonucleotide tggag (t/c) nnntggag (t/c) (where n is an arbitrary nucleotide) that encodes the motif being short having just 15 nucleotides, and the g/c being high, it was extremely difficult to strictly select only those in which
30 the 15 nucleotides have completely hybridized under the usual hybridization conditions.

[0011]

Also, a similar sequence is contained within cDNA encoding proteins other than hemopoietin receptors, starting with various collagens
35 that are thought to be widely distributed and also have high expression amounts, which makes the screening by the above-mentioned plaque

hybridization and RT-PCR highly inefficient.

[0012]

To solve these problems, and to estimate how many different hemopoietic receptor genes actually exist on the human genome, the
5 inventors computer-searched sequences that completely coincided with each probe using all capable oligonucleotide sequences encoding the above-mentioned Trp-Ser-Xaa-Trp-Ser motif as probes.

[0013]

10 Next, among the clones identified by the above search, the nucleotide sequence around the probe sequence of human genome-derived clones (cosmid, BAC, PAC) was converted to the amino acid sequence and compared with the amino acid sequence of known hemopoietin receptors to select genes thought to encode hemopoietin receptor family members.

[0014]

15 From the above search, two clones thought to be hemopoietin receptor genes were identified. One of these was the known GM-CSF β receptor gene (derived from the 22q12.3-13.2 region of chromosome no. 22), and the other (BAC clone AC002303 derived from the 16p12 region of chromosome no. 16) was presumed to encode a novel hemopoietin receptor
20 protein, and this gene was named "NR8."

[0015]

Next, the cDNA thought to encode NR8 was found within the human fetal liver cell cDNA library by RT-PCR using a specific primer designed based on the obtained nucleotide sequence. Furthermore, using this
25 cDNA library as the template, the full-length cDNA NR8 α encoding a transmembrane receptor comprising 361 amino acids was ultimately obtained by 5'-RACE method and 3'-RACE method.

[0016]

30 In the primary structure of NR8 α , a cysteine residue and a proline rich motif conserved between other family members, were well conserved in the extracellular region, and in the intracellular region, the Box 1 motif thought to be involved in signal transduction was well conserved, and therefore, NR8 α was thought to be a typical hemopoietin receptor.

35 [0017]

Furthermore, the inventors revealed the presence of two genes named

NR8 β and NR8 γ as selective splicing products of NR8 α .

[0018]

Therefore, the present invention provides (1) a protein comprising the amino acid sequence from the 1st amino acid Met to the 361st amino acid Ser of SEQ ID NO: 1, or a protein comprising a modified amino acid sequence of said amino acid sequence in which one or more amino acids have been deleted, added and/or substituted with another amino acid and being functionally equivalent to the protein comprising the amino acid sequence from the 1st amino acid Met to the 361st amino acid Ser of SEQ ID NO: 1.

[0019]

The present invention also provides (2) a protein comprising the amino acid sequence from the 1st amino acid Met to the 144th amino acid Leu of SEQ ID NO: 3, or a protein comprising a modified amino acid sequence of said amino acid sequence in which one or more amino acids have been deleted, added and/or substituted with another amino acid and being functionally equivalent to the protein comprising the amino acid sequence from the 1st amino acid Met to the 144th amino acid Leu of SEQ ID NO: 3.

[0020]

The present invention also provides (3) a protein comprising the amino acid sequence from the 1st amino acid Met to the 237th amino acid Ser of SEQ ID NO: 5, or a protein comprising a modified amino acid sequence of said amino acid sequence in which one or more amino acids have been deleted, added and/or substituted with another amino acid and being functionally equivalent to the protein comprising the amino acid sequence from the 1st amino acid Met to the 237th amino acid Ser of SEQ ID NO: 5.

[0021]

The present invention also provides (4) a protein comprising the amino acid sequence from the 1st amino acid Met to the 538th amino acid Ser of SEQ ID NO: 7, or a protein comprising a modified amino acid sequence of said amino acid sequence in which one or more amino acids have been deleted, added and/or substituted with another amino acid and being functionally equivalent to the protein comprising the amino acid sequence from the 1st amino acid Met to the 538th amino acid Ser

of SEQ ID NO: 7.

[0022]

The present invention also provides (5) a protein encoded by a DNA hybridizing to a DNA comprising the nucleotide sequence of SEQ ID NO: 2, said protein being functionally equivalent to a protein comprising the amino acid sequence from the 1st amino acid Met to the 361st amino acid Ser of SEQ ID NO: 1.

[0023]

The present invention also provides (6) a protein encoded by a DNA hybridizing to a DNA comprising the nucleotide sequence of SEQ ID NO: 4, said protein being functionally equivalent to a protein comprising the amino acid sequence from the 1st amino acid Met to the 144th amino acid Leu of SEQ ID NO: 3.

[0024]

The present invention also provides (7) a protein encoded by a DNA hybridizing to a DNA comprising the nucleotide sequence of SEQ ID NO: 6, which is functionally equivalent to a protein comprising the amino acid sequence from the 1st amino acid Met to the 237th amino acid Ser of SEQ ID NO: 5.

[0025]

The present invention also provides (8) a protein encoded by a DNA hybridizing to a DNA comprising the nucleotide sequence of SEQ ID NO: 8, said protein being functionally equivalent to a protein comprising the amino acid sequence from the 1st amino acid Met to the 538th amino acid Ser of SEQ ID NO: 7.

[0026]

The present invention also provides (9) a fusion protein comprising the protein of any one of (1) to (8) and another peptide or polypeptide.

[0027]

The present invention also provides (10) a DNA encoding the protein of any one of (1) to (9).

[0028]

The present invention also provides (11) a vector comprising the DNA of (10).

[0029]

The present invention also provides (12) a transformant harboring

the DNA of (10) in an expressible manner.

[0030]

The present invention also provides (13) a method of producing the protein of any one of (1) to (8), comprising the step of culturing
5 the transformant of (12).

[0031]

The present invention also provides (14) a method of screening a compound that binds to the protein of any one of (1) to (9) comprising the steps of:

10 (a) contacting a test sample with the protein of any one of (1) to (9); and

(b) selecting a compound that comprises an activity to bind to the protein of any one of (1) to (9).

[0032]

15 The present invention also provides (15) an antibody that specifically binds to the protein of any one of (1) to (8).

[0033]

The present invention also provides (16) a method of detecting or measuring the protein of any one of (1) to (9) comprising the steps
20 of contacting a test sample presumed to contain said protein with the antibody of (15), and detecting or measuring the formation of the immune complex between the antibody and the protein.

[0034]

The present invention also provides (17) a DNA specifically
25 hybridizing to a DNA comprising a nucleotide sequence of any one of SEQ ID NOs: 2, 4, 6, and 8 comprising at least 15 nucleotides, and comprising at least 15 nucleotides.

[0035]

[Mode for Carrying Out the Invention]

30 The present invention relates to the novel hemopoietin receptor "NR8." 5'-RACE and 3'-RACE analyses, NR8 genome sequence analysis, and plaque screening analysis revealed the presence of NR8 α , NR8 β , and NR8 γ . The structures of these NR8 genes are shown in Fig. 13. Among the NR8 genes, NR8 β is an alternative splicing product lacking
35 the 5th exon, and can encode two different proteins, a soluble protein in which the CDS ends with a stop codon on the 6th exon that results

from a frame shift following direct coupling to the 4th exon, and a membrane-bound protein lacking the signal sequence starting from the ATG upon the 4th exon.

[0036]

5 Since the soluble protein comprises the same sequence as NR8 α up to the 4th exon, it may function as a soluble receptor. On the other hand, NR8 γ encodes a protein containing a 177 amino acid insertion derived from the NR8 9th intron close to the C terminus of the NR8 α as a result of selective splicing.

10 [0037]

Both NR8 α and NR8 γ encode transmembrane-type hemopoietin receptors. Among the sequences conserved between other hemopoietin receptors that are thought to be involved in signal transduction, a motif resembling Box 1 exists in the intracellular domain of NR8 α and NR8 γ adjacent to the cell membrane. Though low in the degree of conservation, a sequence that is similar to Box 2 also exists, and therefore, NR8 is thought to be a type of receptor in which the signal is transduced by a homodimer.

[0038]

20 The amino acid sequences of the NR8 proteins included in the proteins of the present invention are shown in SEQ ID NO: 1 (NR8 α), SEQ ID NO: 3 (soluble NR8 β), SEQ ID NO: 5 (membrane-bound NR8 β), and SEQ ID NO: 7 (NR8 γ), and the nucleotide sequences of cDNA encoding these proteins are shown in SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, and
25 SEQ ID NO: 8, respectively.

[0039]

Northern blotting for the spleen, thymus, peripheral leucocytes, and lung showed two to three bands in the 5kb and 3 to 4kb regions. Similar sized bands were observed for cell lines HL60 and Raji also, but no expression was seen for other tumor cell lines (HeLa, SW480, A549, G361) and leukemia cell lines (K562, MOLT4).

30 [0040]

The above results suggest that NR8 is specifically expressed on hemopoietic cell lines, especially on granulocytic lines, and B cell lines.

35 [0041]

The above NR8 protein is expected to be applied in medicine. NR8 is expressed in fetal liver, spleen, thymus, and some leukemic cell lines, suggesting the possibility that it might be a receptor of an unknown hemopoietic factor. Therefore, NR8 protein would be a useful material for obtaining this unknown hemopoietic factor.

[0042]

Furthermore, it is possible that NR8 is specifically expressed in limited cell populations within these hemopoietic tissues, and therefore, anti NR8 antibody may be useful as a means of separating these cell populations. Thus separated cell populations can be applied for cell transplant therapy. Anti NR8 antibody is also expected to be applied for the diagnosis and treatment of leukemic diseases represented by leukemia.

[0043]

On the other hand, the soluble protein including the extracellular domain of NR8 protein, or NR8 β , a splicing variant of NR8, may be applied as a decoy-type receptor that is an inhibitor of the NR8 ligand, and is anticipated to be applied in the treatment of diseases in which NR8 is involved, starting with leukemia.

[0044]

The present invention also encompasses a protein that is functionally equivalent to the above-mentioned NR8 protein. Herein "functionally equivalent" means having an equivalent biological activity to the above-mentioned NR8 proteins. Hemopoietic factor receptor protein activity can be given as an example of a biological activity. Such proteins can be obtained by the method of introducing a mutation to the amino acid sequence of a protein. For example, site-specific mutagenesis using a synthetic oligonucleotide primer, can be used to introduce a desired mutation into the amino acid sequence of a protein (Kramer, W. and Fritz, H.J., Methods in Enzymol., 1987, 154, 350-367). This could also be done by a PCR-mediated site-specific mutagenesis system (GIBCO-BRL). Using these methods, the amino acid sequence of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, or SEQ ID NO: 7 can be modified to obtain a protein functionally equivalent to the NR8 protein, in which one or more amino acids in the amino acid sequence of the protein have been deleted, added, and/or substituted by another

amino acid without affecting the biological activity of the protein.

[0045]

As a protein functionally equivalent to the NR8 protein of the invention, the following are given: one in which one or two or more, preferably, two to 30, more preferably, two to ten amino acids are deleted in any one of the amino acid sequences of SEQ ID NO: 1, SEQ ID NO: 3, or SEQ ID NO: 5; one in which one or two or more, preferably, two to 30, more preferably, two to ten amino acids have been added into any one of the amino acid sequences of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, and SEQ ID NO: 7; or one in which one or two or more, preferably, two to 30, more preferably, two to ten amino acids have been substituted with other amino acids in any one of the amino acid sequences of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, or SEQ ID NO: 7.

[0046]

It is already known that a protein comprising a modified amino acid sequence of a certain amino acid sequence in which one or more amino acid residues have been deleted, added, and/or substituted with another amino acid, still maintains its biological activity (Mark, D. F. et al., Proc. Natl. Acad. Sci. USA, 1984, 81, 5662-5666; Zoller, M. J. & Smith, M., Nucleic Acids Research, 1982, 10, 6487-6500; Wang, A. et al., Science, 224, 1431-1433; Dalbadie-McFarland, G. et al., Proc. Natl. Acad. Sci. USA, 1982, 79, 6409-6413).

[0047]

For example, a fusion protein can be given as a protein in which one or more amino acid residues have been added to the NR8 protein of the present invention. A fusion protein is made by fusing the NR8 protein of the present invention with another peptide or protein and is encompassed in the present invention. A fusion protein can be prepared by ligating DNA encoding the NR8 protein of the present invention with DNA encoding another peptide or protein so as the frames match, introducing this into an expression vector, and expressing the fusion gene in a host. Methods commonly known can be used for preparing such a fusion gene. There is no restriction as to the other peptide or protein that is fused to the protein of this invention.

[0048]

For example, FLAG (Hopp, T.P. et al., Biotechnology, 1988, 6, 1204-1210), 6x His constituting six histidine (His) residues, 10x His, Influenza agglutinin (HA), human c-myc fragment, VSV-GP fragment, p18HIV fragment, T7-tag, HSV-tag, E-tag, SV40T antigen fragment, lck tag, α -tubulin fragment, B-tag, Protein C fragment, and such well-known peptides can be used. Examples of proteins are, glutathione-S-transferase (GST), Influenza agglutinin (HA), immunoglobulin constant region, β -galactosidase, maltose-binding protein (MBP), etc. Commercially available DNAs encoding these may also be used to prepare fusion proteins.

[0049]

The protein of the invention can also be encoded by a DNA that hybridizes under stringent conditions to a DNA comprising any one of the nucleotide sequences of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, and SEQ ID NO: 8. Such a protein also includes a protein functionally equivalent to the above-mentioned NR8 protein. Stringent conditions can be suitably selected by one skilled in the art, and for example, low-stringent conditions can be given. Low-stringent conditions are, for example, 42°C, 2x SSC, and 0.1% SDS, and preferably, 50°C, 2x SSC, and 0.1% SDS. More preferable are highly stringent conditions, for example, 65°C, 2x SSC, and 0.1% SDS. Under these conditions, the higher the temperature is raised, the higher the homology of the obtained DNA will be.

[0050]

The present invention also includes a protein that is functionally equivalent to the above NR8 protein, which has also a homology with a protein comprising any one of the amino acid sequences of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, and SEQ ID NO: 7. A protein having a homology means, a protein having at least 70%, preferably at least 80%, more preferably at least 90%, even more preferably, at least 95% homology to any one of the amino acid sequences of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, and SEQ ID NO: 7. The homology of a protein can be determined by the algorithm in "Wilbur, W.J. and Lipman, D.J. Proc. Natl. Acad. Sci. USA, 1983, 80, 726-730."

[0051]

In the protein of the invention, the amino acid sequence, molecular

weight, isoelectric point, the presence or absence of the sugar chain, and its form differ according to the producing cells, host, or purification method described below. However, as long as the obtained protein comprises a hemopoietic factor receptor protein activity, it is included in the present invention.

[0052]

For example, if the protein of the present invention is expressed in prokaryotic cells such as *E. coli*, a methionine residue is added at the N-terminus of the amino acid sequence of the expressed protein.

If the protein of the present invention is expressed in eukaryotic cells such as mammalian cells, the N-terminal signal sequence is removed. The protein of the present invention includes these proteins.

[0053]

For example, as a result of analyzing the protein of the invention based on the method in "Von Heijne, G., Nucleic Acids Research, 1986, 14, 4683-4690," it was presumed that the signal sequence is from the 1st Met to the 19th Gly in the amino acid sequence of SEQ ID NO: 1. Therefore, the present invention encompasses a protein comprising the sequence from the 20th Cys to 361st Ser in the amino acid sequence of SEQ ID NO: 1.

[0054]

To produce the protein of the invention, the obtained DNA is incorporated into an expression vector in a manner that the DNA is expressible under the regulation of an expression regulatory region, for example, an enhancer or promoter. Next, host cells are transformed by this expression vector to express the protein.

[0055]

Specifically, the protein can be produced as follows. When mammalian cells are used, DNA comprising a commonly used useful promoter/enhancer, DNA encoding the protein of the invention, and the poly A signal that is functionally bound to the 3' side downstream of the protein-encoding DNA, or a vector containing it, is constructed. For example, as the promoter/enhancer, human cytomegalovirus immediate early promoter/enhancer can be given.

[0056]

Also, as other promoters/enhancers that can be used for protein

expression, viral promoters/enhancers of retroviruses, polyomaviruses, adenoviruses, simian virus 40 (SV40), and such, and promoters/enhancers derived from mammalian cells, such as that of human elongation factor 1 α (HEF1 α) can be used.

5 [0057]

For example, a protein can be easily expressed by following the method of Mulligan et al. (Nature, 1979, 277, 108) when using the SV40 promoter/enhancer, and the method of Mizushima et al. (Nucleic Acids Res., 1990, 18, 5322) when using the HEF1 α promoter/enhancer.

10 [0058]

When using *E. coli*, well-used useful promoters, the signal sequence for polypeptide secretion, and genes to be expressed, may be functionally bound to express the desired gene. For example, lacZ promoter and araB promoter may be used as promoters. When using the lacZ promoter, the method of Ward et al. (Nature, 1098, 341, 544-546; FASEB J., 1992, 6, 2422-2427), and when using the araB promoter, the method of Better et al. (Science, 1988, 240, 1041-1043) may be followed.

[0059]

When producing the protein into the periplasm of *E. coli*, the pelB (Lei, S. P. et al., J. Bacteriol., 1987, 169, 4379) signal sequence may be used as a protein secretion signal.

[0060]

A replication origin derived from SV40, polyomavirus, adenovirus, bovine papillomavirus (BPV), and such may be used. To amplify gene copies in host cell lines, the expression vector may include an aminoglycoside transferase (APH) gene, thymidine kinase (TK) gene, *E. coli* xanthine guanine phosphoribosyl transferase (Ecogpt) gene, dihydrofolate reductase (dhfr) gene, and such as a selective marker.

[0061]

The expression vector used to produce the protein of the invention may be any, as long as it's an expression vector that is suitably used for the present invention. Mammalian expression vectors, for example, pEF and pCDM8; insect-derived expression vectors, for example, pBacPAK8; plant-derived expression vectors, for example, pMH1 and pMH2; animal virus-derived expression vectors, for example, pHSV, pMV, and pAdexLcw; retrovirus-derived expression vectors, for example,

pZIpneo; yeast-derived expression vectors, for example, pNV11 and SP-Q01; *Bacillus subtilis*-derived expression vectors, for example, pPL608 and pKTH50; *E. coli*-derived expression vectors, for example, pQE, pGEAPP, pGEMEAPP, and pMALp2 can be given as expression vectors of this invention.

[0062]

Not only vectors of the present invention that produce the protein of the invention *in vivo* and *in vitro*, but also those that are used for gene therapy of mammals, for example humans.

[0063]

When introducing the expression vector of the present invention constructed above into a host cell, well-known methods, for example the calcium phosphate method (Virology, 1973, 52, 456-467), electroporation (EMBO J., 1982, 1, 841-845), and such may be used.

[0064]

In the present invention, an arbitrary production system may be used to produce the protein. *In vitro* and *in vivo* production systems are known as production systems for producing proteins. Production systems using eukaryotic cells and prokaryotic cells may be used as *in vitro* production systems.

[0065]

When using eukaryotic cells, production systems using, for example, animal cells, plant cells, and fungal cells are known. As animal cells used, for example, mammalian cells such as CHO (J. Exp. Med., 1995, 108, 945), COS, myeloma, baby hamster kidney (BHK), HeLa, or Vero, amphibian cells such as *Xenopus* oocytes (Valle, et al., Nature, 1981, 291, 358-340), insect cells such as sf9, sf21, or Tn5, are known. As CHO cells, especially DHFR gene-deficient CHO cell, dhfr-CHO (Proc. Natl. Acad. Sci. USA, 1980, 77, 4216-4220), and CHO K-1 (Proc. Natl. Acad. Sci. USA, 1968, 60, 1275) can be suitably used.

[0066]

Nicotiana tabacum-derived cells are well known as plant cells, and these can be callus cultured. As fungal cells, yeasts such as the *Saccharomyces* genus, for example, *Saccharomyces cerevisiae*, filamentous bacteria such as the *Aspergillus* genus, for example, *Aspergillus niger* are known.

[0067]

Bacterial cells may be used as prokaryotic production systems. As bacterial cells, *E. coli* and *Bacillus subtilis* are known.

[0068]

5 Proteins can be obtained by transforming these cells with the objective DNA, and culturing the transformed cells *in vitro* according to well-known methods. For example, DMEM, MEM, RPMI1640, and IMDM can be used as culture media. At that instance, fetal calf serum (FCS) and such serum supplements may be added in the above media, or a
10 serum-free culture medium may be used. The pH is preferably about 6 to 8. Culture is usually done at about 30°C to 40°C, for about 15 to 200 hr, and medium changes, aeration, and stirring are done as necessary.

[0069]

15 On the other hand, production systems using animals and plants may be given as *in vivo* production systems. The objective gene is introduced into the plant or animal, and the protein is produced within the plant or animal, and recovered. "Host" as used in the present invention encompasses such animals and plants as well.

20 [0070]

When using animals, mammalian and insect production systems can be used. As mammals, goats, pigs, sheep, mice, and cattle may be used (Vicki Glaser, SPECTRUM Biotechnology Applications, 1993). Transgenic animals may also be used when using mammals.

25 [0071]

For example, the objective DNA is inserted within a gene encoding a protein produced intrinsically into milk, such as goat β casein, to prepare a fusion gene. The DNA fragment containing the fusion gene is injected into a goat's embryo, and this embryo is implanted in
30 a female goat. The protein is collected from the milk of the transgenic goats produced from the goat that received the embryo, and descendents thereof. To increase the amount of protein-containing milk produced from the transgenic goat, a suitable hormone/hormones may be given to the transgenic goats (Ebert, K.M. et al., Bio/Technology, 1994,
35 12, 699-702).

[0072]

Silk worms may be used as insects. When using the silk worm, it is infected with a baculovirus to which the objective DNA has been inserted, and the desired protein is obtained from the body fluids of the silk worm (Susumu, M. et al., Nature, 1985, 315, 592-594).

5 [0073]

When using plants, for example, tobacco can be used. In the case of tobacco, the objective DNA is inserted into a plant expression vector, for example pMON 530, and this vector is introduced into a bacterium such as *Agrobacterium tumefaciens*. This bacterium is
10 infected to tobacco, for example *Nicotiana tabacum*, to obtain the desired polypeptide from tobacco leaves (Julian, K.-C. Ma et al., Eur. J. Immunol., 1994, 24, 131-138).

[0074]

The thus-obtained protein of the invention is isolated from within
15 and without cells, or from hosts, and can be purified as a substantially pure homogenous protein. The separation and purification of the protein is not limited to any specific method and can be done using ordinary separation and purification methods used to purify proteins. For example, chromatography, filtration, ultrafiltration, salting
20 out, solvent precipitation, solvent extraction, distillation, immunoprecipitation, SDS-polyacrylamide gel electrophoresis, isoelectric focusing, dialysis, recrystallization, and such may be suitably selected, or combined to separate/purify the protein.

[0075]

25 As chromatographies, for example, affinity chromatography, ion exchange chromatography, hydrophobic chromatography, gel filtration, reversed-phase chromatography, adsorption chromatography, and such can be exemplified (Strategies for Protein Purification and Characterization: A Laboratory Course Manual. Ed Daniel R. Marshak
30 et al., Cold Spring Harbor Laboratory Press, 1996). These chromatographies can be done by liquid chromatography such as HPLC, FPLC, and the like. The present invention encompasses proteins highly purified by using such purification methods.

[0076]

35 Proteins can be arbitrarily modified, or peptides may be partially excised by treating the proteins with appropriate modification enzymes

prior to or after the purification. Trypsin, chymotrypsin, lysyl endopeptidase, protein kinase, glucosidase, and such are used as protein modification enzymes.

[0077]

5 The present invention includes a partial peptide comprising the active center of a protein comprising any one of the amino acid sequences of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, and SEQ ID NO: 7. A partial peptide of the protein of the present invention is, for example, a partial peptide of the molecules of the protein, which contains one
10 or more regions of the hydrophilic region and hydrophobic region presumed by hydrophobicity plot analysis. These partial peptides may contain the whole hydrophilic region or a part of it, and may contain the whole hydrophobic region or a part of it. For example, soluble proteins and proteins comprising extracellular regions of the protein
15 of the invention, are also encompassed in the invention.

[0078]

The partial peptides of the protein of the invention may be produced by genetic engineering techniques, well-known peptide synthesizing methods, or by excising the protein of the invention by a suitable
20 peptidase. As peptide synthesizing methods, the solid-phase synthesizing method, and the liquid-phase synthesizing method may be used.

[0079]

The present invention also relates to a DNA encoding the protein
25 of the invention. A cDNA encoding the protein of the invention may be obtained by, for example, screening a human cDNA library using the probe described herein.

[0080]

Using the obtained cDNA or cDNA fragment as a probe, cDNA can also
30 be obtained from other cells, tissues, organs, or species by further screening cDNA libraries. cDNA libraries may be prepared by, for example, the method of Sambrook, J. et al., Molecular Cloning, Cold Spring Harbor Laboratory Press (1989), or commercially available cDNA libraries may be used.

35 [0081]

By determining the nucleotide sequence of the obtained cDNA, the

translation region encoded by it can be determined, and the amino acid sequence of the protein of the present invention can be obtained. Furthermore, genomic DNA can be isolated by screening the genomic DNA library using the obtained cDNA as a probe.

5 [0082]

Specifically, this can be done as follows. First, mRNA is isolated from cells, tissues, and organs expressing the protein of the invention. For this mRNA isolation, whole RNA is prepared using well-known methods, for example, guanidine ultracentrifugation method (Chirgwin, J.M. et al., Biochemistry, 1979, 18, 5294-5299), the AGPC method (Chomczynski, P. and Sacchi, N., Anal. Biochem., 1987, 162, 156-159), and such, and purified using the mRNA Purification Kit (Pharmacia), etc. mRNA may be directly prepared using the QuickPrep mRNA Purification Kit (Pharmacia).

15 [0083]

cDNA is synthesized using reverse transcriptase from the obtained mRNA. cDNA can be synthesized using the AMV Reverse Transcriptase First-strand cDNA Synthesis Kit (SEIKAGAKU CORPORATION), etc. Also, cDNA synthesis and amplification may also be done using the probe described herein by following the 5'-RACE method (Frohman, M.A. et al., Proc. Natl. Acad. Sci. U.S.A., 1988, 85, 8998-9002; Belyavsky, A. et al., Nucleic Acids Res., 1989, 17, 2919-2932) using the polymerase chain reaction (PCR) and the 5'-Ampli FINDER RACE KIT (Clontech).

[0084]

25 The objective DNA fragment is prepared from the obtained PCR product and ligated with vector DNA. Thus, a recombination vector is created, introduced into *E.coli*, etc. and colonies are selected to prepare the desired recombination vector. The nucleotide sequence of the objective DNA may be verified by known methods, for example, the dideoxy
30 nucleotide chain termination method.

[0085]

In the DNA of the invention, a sequence with a higher expression efficiency can be designed by considering the codon usage frequency of hosts used for the expression (Grantham, R. et al., Nucleic Acids
35 Research, 1981, 9, p43-p74). The DNA of the invention may also be modified using commercially available kits and known methods. For

example, digestion by restriction enzymes, insertion of synthetic oligonucleotides and suitable DNA fragments, addition of linkers, insertion of a start codon (ATG) and/or stop codon (ATT, TGA, or TAG), and such can be given.

5 [0086]

The DNA of the present invention encompasses DNA comprising the nucleotide sequence from the 441st nucleotide A to the 1523rd nucleotide C in the nucleotide sequence of SEQ ID NO: 2, DNA comprising the nucleotide sequence from the 441st nucleotide A to the 872nd nucleotide A in the nucleotide sequence of SEQ ID NO: 4, DNA comprising the nucleotide sequence from the 659th nucleotide A to the 1368th nucleotide C in the nucleotide sequence of SEQ ID NO: 6, and DNA comprising the nucleotide sequence from the 441st nucleotide A to the 2054th nucleotide C in the nucleotide sequence of SEQ ID NO: 8.

15 [0087]

The DNA of the present invention encompasses DNA that hybridizes under stringent conditions to the DNA comprising any one of the nucleotide sequences of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, and SEQ ID NO: 8, which also includes a DNA encoding a protein functionally equivalent to the NR8 protein.

20 [0088]

Stringent conditions can be suitably selected by one skilled in the art, and for example, low-stringent conditions can be given. Low-stringent conditions are, for example, 42°C, 2x SSC, and 0.1% SDS, and preferably 50°C, 2x SSC, and 0.1% SDS. More preferable are highly stringent conditions, for example, 65°C, 2x SSC, and 0.1% SDS. Under these conditions, the higher the temperature is raised, the higher the homology of the obtained DNA will be. The above DNA is preferably natural DNA such as cDNA and chromosomal DNA.

30 [0089]

As shown in Examples, the mRNA of the gene hybridizing to cDNA encoding the protein of the invention was distributed in various human tissues. Therefore, the above-mentioned natural DNA may be, for example, genomic DNA and cDNA derived from tissues in which the mRNA that hybridizes to the cDNA encoding the protein of the invention is detected in Examples. The DNA encoding the protein of the invention

may be cDNA, genomic DNA, or synthetic DNA.

[0090]

The protein of the invention is useful in screening a compound that binds to it. Namely, the protein of the invention is used in the screening method that comprises the steps of contacting a test sample expected to contain a compound that binds to the protein of the invention with the protein of the invention, and selecting the compound that comprises an activity to bind to the protein of the invention.

[0091]

As methods for screening a compound that comprises an activity to bind to the protein of the invention, numerous methods usually used by those skilled in the art can be employed. The protein of the invention that is used for the screening of the invention may be a recombinant, natural, or partial peptide. A compound comprising an activity to bind to the protein of the invention may be a protein comprising a binding activity, or it may be a chemically synthesized compound having a binding activity.

[0092]

As a test sample that is used in the screening method of the present invention, for example, peptides, purified or crudely purified proteins, non-peptide compounds, synthetic compounds, microbial fermentation products, extracts of marine organisms, plant extracts, cell extracts, animal tissue extracts, and such can be given. These test samples may be novel compounds, or well-known compounds.

[0093]

A protein that binds to the protein of the invention can be screened by, for example, using the West-western blotting method (Skolnik, E.Y. et al., Cell, 1991, 65, 83-90). cDNA is isolated from cells, tissues, and organs presumed to express the protein binding to the protein of the invention, this is inserted into phage vectors, for example, λ gt11, ZAPII, and such, to make a cDNA library, expressed on a plate containing a culture medium, the proteins expressed are fixed on a filter, this filter is reacted with the labeled, purified protein of the invention, and plaques expressing the protein bound to the protein of the invention are detected by the labels. As methods

to label the protein of the invention, the method that uses the binding ability of avidin and biotin, the method of using an antibody that specifically binds to the protein of the invention or the peptide or polypeptide fused to the protein of the invention, the method of
5 using radioisotopes, or fluorescence, and such can be given.

[0094]

A ligand that binds specifically to the protein of the invention can be screened by, preparing a chimeric receptor by ligating the extracellular domain of the protein of the invention with the
10 intracellular domain containing the transmembrane domain of a hemopoietin receptor protein comprising a known signal transduction ability, expressing this chimeric receptor on the cell surface of a suitable cell line, preferably, a cell line that can survive and proliferate under the presence of a suitable growth factor (a growth
15 factor-dependent cell line), and culturing the cell line by adding a material that is expected to contain various growth factors, cytokines, or hemopoietic factors. This method uses the fact that the above-mentioned growth factor-dependent cell line survives and proliferates only when a ligand that specifically binds to the
20 extracellular domain of the protein of the invention exists within the test material. Known hemopoietic receptors are, for example, the thrombopoietin receptor, erythropoietin receptor, G-CSF receptor, gp130, etc. However, the partner of the chimeric receptor used in the screening of the invention is not limited to these known hemopoietic
25 receptors, and any may be used as long as a structure needed for the signal transduction activity is contained in the cytoplasmic domain. Growth factor-dependent cell lines are for example, IL-3-dependent cell lines starting with BaF3 and FDC-P1.

[0095]

As a ligand that specifically binds to the protein of the invention, the possibility of not only soluble proteins, but also cell
30 membrane-binding proteins can be envisaged, though rare. In such cases, screening can be done by labeling the protein containing only the extracellular domain of the protein of the invention, or a fusion
35 protein in which the partial sequence of another soluble protein has been added to this extracellular domain, and measuring the binding

with cells expected to express the ligand. As examples of proteins containing only the extracellular domain of the protein of the invention, for example, a soluble receptor protein artificially made by inserting a stop codon to the N terminal side of the transmembrane domain, or NR8 β soluble protein may be used. On the other hand, as a fusion protein in which the partial sequence of another soluble protein has been added to the extracellular domain of the protein of the invention, for example, proteins prepared by adding immunoglobulin Fc site, FLAG peptide, etc. to the C terminus of the extracellular domain can be used. These soluble labeled proteins can be used in the detection in the above-described West-western blotting method.

[0096]

A protein that binds to the protein of the invention can be screened by using the two-hybrid system (Fields, S. and Sternglanz, R., Trends. Genet., 1994, 10, 286-292).

[0097]

In the two-hybrid system, an expression vector containing DNA encoding the fusion protein between the protein of the invention and one subunit of a heterodimeric transcriptional regulatory factor, and an expression vector containing DNA made by ligating DNA encoding the other subunit of the heterodimeric transcriptional regulatory factor and a desired cDNA used as a test sample are introduced into cells and expressed. If the protein encoded by the cDNA binds with the protein of the invention and the transcriptional regulatory factor forms a heterodimer, a reporter gene constructed in the cell beforehand will be expressed. Therefore, a protein binding to the protein of the invention can be selected by detecting or measuring the expression level of the reporter gene.

[0098]

Specifically, the DNA encoding the protein of the invention and the gene encoding the DNA binding domain of LexA are ligated so as the frames match to prepare an expression vector. Next, the desired cDNA and the gene encoding GAL4 transcription activation domain are ligated to prepare an expression vector.

[0099]

Cells into which the HIS3 gene has been incorporated (the

transcription of HIS3 gene is regulated by the promoter having a LexA binding motif) are transformed by the above two-hybrid system expression plasmids, and then incubated on a histidine-free synthetic culture medium. Herein, cells only grow when a protein interaction is present. Thus, the increase in reporter gene expression can be examined by the growth rate of the transformant.

[0100]

Other than the HIS3 gene, for example, the luciferase gene, plasminogen activator inhibitor type1 (PAI-1) gene and such can be used as reporter genes.

[0101]

The two-hybrid system may be constructed according to the usual methods, or a commercially available kit may be used. As commercially available two-hybrid system kits, the MATCHMARKER Two-Hybrid System, Mammalian MATCHMARKER Two-Hybrid Assay Kit (both by CLONTEC), and HybriZAP Two-Hybrid Vector System (Stratagene) can be given.

[0102]

A protein binding to the protein of the invention can be screened by affinity chromatography. Namely, the protein of the invention is immobilized onto a carrier of an affinity column, and a test sample presumed to express a protein binding to the protein of the invention is applied to the column. As this test sample, a cell culture supernatant, cell extract, cell lysate, and such may be used. After applying the test sample, the column is washed to obtain the protein binding to the protein of the invention.

[0103]

The compound isolated by the screening method of the invention is a candidate drug for promoting or inhibiting the activity of the protein of the invention. The compound obtained by using the screening method of the invention encompasses a compound resulting from modifying the compound having an activity to bind to the protein of the invention by adding, deleting, and/or replacing a part of the structure.

[0104]

When using the compound obtained by the screening method of the invention as drugs for humans and mammals such as, mice, rats, guinea pigs, rabbits, chicken, cats, dogs, sheep, pigs, cattle, monkeys,

sacred baboons, and chimpanzees, the drug may be administered using ordinary means.

[0105]

For example, according to the need, the drugs can be taken orally
5 as sugar-coated tablets, capsules, elixirs, and microcapsules, or
parenterally in the form of injections of sterile solutions or
suspensions with water or any other pharmaceutically acceptable liquid.
For example, the compounds comprising the activity to bind to the
protein of the invention can be mixed with physiologically acceptable
10 carriers, flavoring agents, excipients, vehicles, preservatives,
stabilizers, and binders, in a unit dose form required for generally
accepted drug implementation. The amount of active ingredients in
these preparations makes a suitable dosage within the indicated range
acquirable.

15 [0106]

Examples of additives that can be mixed to tablets and capsules
are, binders such as gelatin, corn starch, tragacanth gum, and arabic
gum; excipients such as crystalline cellulose; swelling agents such
as cornstarch, gelatin, and alginic acid; lubricants such as magnesium
20 stearate; sweeteners such as sucrose, lactose, or saccharin; and
flavoring agents such as peppermint, Gaultheria adeno-thrix oil, and
cherry. When the unit dosage form is a capsule, a liquid carrier,
such as oil, can also be included in the above additives. Sterile
compositions for injections can be formulated following usual drug
25 implementations using vehicles such as distilled water used for
injections.

[0107]

For example, physiological saline and isotonic liquids including
glucose or other adjuvants, such as D-sorbitol, D-mannose, D-mannitol,
30 and sodium chloride, can be used as aqueous solutions for injections.
These can be used in conjunction with suitable solubilizers, such
as alcohol, specifically ethanol, polyalcohols such as propylene
glycol and polyethylene glycol, non-ionic surfactants, such as
Polysorbate 80 (TM) and HCO-50.

35 [0108]

Sesame oil or soy-bean oil can be used as a oleaginous liquid

and may be used in conjunction with benzyl benzoate or benzyl alcohol as a solubilizer; may be formulated with a buffer such as phosphate buffer and sodium acetate buffer; a pain-killer such as procaine hydrochloride; a stabilizer such as benzyl alcohol and phenol; and an anti-oxidant. The prepared injection is usually filled into a suitable ampule.

[0109]

Although the dosage of the compound that has the activity to bind to the protein of the invention varies according to symptoms, the daily dose is generally about 0.1 to about 100 mg, preferably about 1.0 to about 50 mg, and more preferably about 1.0 to about 20 mg, when administered orally to an adult (body weight 60 kg).

[0110]

When given parenterally, the dose differs according to the patient, target organ, symptoms, and method of administration, but the daily dose is usually about 0.01 to about 30 mg, preferably about 0.1 to about 20 mg and more preferably about 0.1 to about 10 mg for an adult (body weight 60 kg) when given as an intravenous injection. Also, in the case of other animals too, it is possible to administer an amount converted to 60 kg of body-weight.

[0111]

The antibody of the present invention can be obtained as a monoclonal antibody or a polyclonal antibody using well-known methods.

[0112]

The antibody that specifically binds to the protein of the invention can be prepared by using the protein of the invention as a sensitizing antigen for immunization according to usual immunizing methods, fusing the obtained immunized cells with known parent cells by ordinary cell fusion methods, and screening for antibody producing cells using the usual screening techniques.

[0113]

Specifically, a monoclonal or polyclonal antibody that binds to the proteins of the invention may be prepared as follows.

[0114]

For example, the protein of the invention that is used as a sensitizing antigen for obtaining the antibody is not restricted by

the animal species from which it is derived, but is preferably a protein derived from mammals, for example, humans, mice, or rats, especially from humans. Proteins of human origin can be obtained by using the nucleotide sequence or amino acid sequence disclosed herein.

5 [0115]

The protein that is used as a sensitizing antigen in the present invention can be a protein that comprises the biological activity of all the proteins described herein. Partial peptides of the proteins may also be used. As partial peptides of the proteins, for example,
10 the amino (N) terminal fragment of the protein, and the carboxy (C) terminal fragment can be given. "Antibody" as used herein means an antibody that specifically reacts with the full-length or fragment of the protein.

[0116]

15 A gene encoding the protein of the invention or a fragment thereof is inserted into a well-known expression vector, and after transforming the host cells described herein, the objective protein or a fragment thereof is obtained from within and without the host cell, or from the host using well-known methods, and this protein can be used as
20 a sensitizing antigen. Also, cells expressing the protein, cell lysates, or chemically synthesized protein of the invention may be used as a sensitizing antigen.

[0117]

The mammals that are immunized by the sensitizing antigen are not
25 restricted, but it is preferable to select the animal by considering the adaptability with the parent cells used in cell fusion. Generally, an animal belonging to Rodentia, Lagomorpha, or Primates is used.

[0118]

As animals belonging to Rodentia, for example, mice, rats, hamsters,
30 and such are used. As animals belonging to Lagomorpha, for example rabbits, as Primates, for example monkeys, are used. As monkeys, monkeys of the infraorder Catarrhini (Old World Monkeys), for example, cynomolgus monkeys, rhesus monkeys, sacred baboons, chimpanzees, etc., are used.

35 [0119]

To immunize animals with the sensitizing antigen, well-known

methods may be used. For example, the sensitizing antigen is generally injected into mammals intraperitoneally or subcutaneously. Specifically, the sensitizing antigen is suitably diluted, suspended in physiological saline or phosphate-buffered saline (PBS), mixed with a suitable amount of a general adjuvant if desired, for example, with Freund's complete adjuvant, emulsified and injected into the mammal. Thereafter, the sensitizing antigen suitably mixed with Freund's incomplete adjuvant is preferably given several times every four to 21 days. A suitable carrier can also be used when immunizing an animal with the sensitizing antigen. After the immunization, the elevation in the serum antibody level is detected by usual methods.

[0120]

Polyclonal antibodies against the protein of the invention can be obtained as follows. After verifying that the desired serum antibody level has been reached, blood is withdrawn from the mammal sensitized with the antigen. Serum is isolated from this blood using well-known methods. The serum containing the polyclonal antibody may be used as the polyclonal antibody, or according to needs, the polyclonal antibody-containing fraction may be further isolated from the serum.

[0121]

To obtain monoclonal antibodies, after verifying that the desired serum antibody level has been reached in the mammal sensitized with the above-described antigen, immunocytes are taken from the mammal and used for cell fusion. At this instance, immunocytes that are preferably used for cell fusion are splenocytes. As parent cells fused with the above immunocytes, preferable are mammalian myeloma cells, more preferable are, myeloma cells that have attained the feature of distinguishing fusion cells by agents.

[0122]

For the cell fusion between the above immunocytes and myeloma cells, for example, the method of Milstein et al. (Galfre, G. and Milstein, C., Methods Enzymol., 1981, 73, 3-46) is basically well known.

[0123]

The hybridoma obtained from cell fusion is selected by culturing in a usual selective culture medium, for example, HAT culture medium (hypoxanthine, aminopterin, thymidine-containing culture medium).

The culture in this HAT medium is continued for a period sufficient enough for cells (non-fusion cells) other than the objective hybridoma to perish, usually from a few days to a few weeks. Next, the usual limiting dilution method is carried out, and the hybridoma producing the objective antibody is screened and cloned.

[0124]

Other than the above method of obtaining a hybridoma by immunizing an animal other than humans with the antigen, a hybridoma producing the objective human antibodies comprising the activity to bind to proteins can be obtained by the method of sensitizing human lymphocytes, for example, human lymphocytes infected with the EB virus, with proteins, protein-expressing cells, or lysates thereof *in vitro*, fusing the sensitized lymphocytes with myeloma cells derived from human, for example U266, having the capacity of permanent cell division (Unexamined Published Japanese Patent Application (JP-A) No. Sho 63-17688).

[0125]

Moreover, human antibody against the protein can be obtained using a hybridoma made by fusing myeloma cells with antibody-producing cells obtained by immunizing a transgenic animal comprising a repertoire of human antibody genes with an antigen such as a protein, protein-expressing cells, or a cell lysate thereof WO92/03918, WO93/2227, WO94/02602, WO94/25585, WO96/33735, and WO96/34096).

[0126]

Other than producing antibodies by using hybridoma, antibody-producing immunocytes such as sensitized lymphocytes that are immortalized by oncogenes may also be used.

[0127]

Such monoclonal antibodies can also be obtained as recombinant antibodies produced by using the genetic engineering technique (for example, Borrebaeck, C.A.K. and Larrick, J.W., THERAPEUTIC MONOCLONAL ANTIBODIES, Published in the United Kingdom by MACMILLAN PUBLISHERS LTD, 1990). Recombinant antibodies are produced by cloning the encoding DNA from immunocytes such as hybridoma or antibody-producing sensitized lymphocytes, incorporating this into a suitable vector, and introducing this vector into a host to produce the antibody. The

present invention encompasses such recombinant antibodies as well.

[0128]

The antibody of the present invention may be an antibody fragment or a modified-antibody as long as it binds to the protein of the invention.

5 For example, Fab, F(ab')₂, Fv, or single chain Fv in which the H chain Fv and the L chain Fv are suitably linked by a linker (scFv, Huston, J.S. et al., Proc. Natl. Acad. Sci. U.S.A., 1988, 85, 5879-5883) can be given as antibody fragments. Specifically, antibody fragments are produced by treating an antibody with an enzyme, for example, papain,
10 pepsin, etc. or by constructing a gene encoding an antibody fragment, introducing this into an expression vector, and expressing this vector on suitable host cells (for example, Co, M.S. et al., J. Immunol., 1994, 152, 2968-2976; Better, M. and Horwitz, A.H., Methods Enzymol., 1989, 178, 476-496; Pluckthun, A. and Skerra, A., Methods Enzymol.,
15 1989, 178, 497-515; Lamoyi, E., Methods Enzymol., 1986, 121, 652-663; Rousseaux, J. et al., Methods Enzymol., 1986, 121, 663-669; Bird, R.E. and Walker, B.W., Trends Biotechnol., 1991, 9, 132-137).

[0129]

As a modified antibody, an antibody bound to various molecules
20 such as polyethylene glycol (PEG) can be used. The present antibody encompasses such modified antibodies as well. To obtain such a modified antibody, chemical modifications are done to the obtained antibody. These methods are already established in the field.

[0130]

25 The antibody of the invention may be obtained as a chimeric antibody comprising non-human antibody-derived variable region and a human antibody-derived constant region, or as a humanized antibody comprising non-human antibody-derived complementarity determining region (CDR), and human antibody-derived framework region (FR) and
30 a constant region.

[0131]

Antibodies thus obtained can be purified till uniform. The separation and purification methods for separating and purifying the antibody used in the present invention may be any method usually used
35 for proteins, and is not in the least limited. Antibody concentration of the above mentioned antibody can be assayed by measuring the

absorbance, or by the enzyme-linked immunosorbent assay (ELISA), etc.

[0132]

Also, as methods that assay the antigen-binding activity of the antibody of the invention, ELISA, enzyme immunoassay (EIA), radio
5 immunoassay (RIA), or fluorescent antibody method can be given. For example, when using ELISA, the protein of the invention is added to a plate coated with the antibody of the invention, and next, the objective antibody sample, for example, culture supernatants of antibody-producing cells, or purified antibodies are added. Then,
10 secondary antibody recognizing the antibody, which is labeled by alkaline phosphatase and such enzymes, is added, the plate is incubated and washed, and absorbance is measured to evaluate the antigen-binding activity after adding an enzyme substrate such as p-nitrophenyl phosphate. As the protein, a protein fragment, for example, a fragment
15 comprising a C terminus, or a fragment comprising an N terminus may be used. To evaluate the activity of the antibody of the invention, BIAcore (Pharmacia) may be used.

[0133]

By using these methods, the antibody of the invention and a sample
20 presumed to contain the protein of the invention are contacted, and the protein of the invention is detected or assayed by detecting or assaying the immune complex of the above-mentioned antibody and protein.

[0134]

25 A method of detecting or assaying the protein of the invention is useful in various experiments using proteins as it can specifically detect or assay the proteins.

[0135]

The present invention also encompasses a DNA specifically
30 hybridizing to a DNA comprising a nucleotide sequence of any one of SEQ ID NOs: 2, 4, 6, and 8, or its complementary DNA, and comprising at least 15 nucleotides. Namely, a probe that can selectively hybridize to the DNA encoding the protein of the invention, or a DNA complementary to the above DNA, a nucleotide or nucleotide derivative, for example,
35 antisense oligonucleotide, ribozyme, and such are included.

[0136]

The present invention also encompasses an antisense oligonucleotide that hybridizes to any portion of any one of the nucleotide sequences shown in, for example, SEQ ID NOs: 2, 4, 6, and 8. This antisense oligonucleotide is preferably one against at least
5 15 continuous nucleotides in any one of the nucleotide sequences of SEQ ID NOs: 2, 4, 6, and 8. More preferable is the above-mentioned antisense oligonucleotide against the above-mentioned at least 15 continuous nucleotides containing a translation start codon.

[0137]

10 Derivatives or modified products of antisense oligonucleotides can be used as antisense oligonucleotides. As such modified products, for example, lower alkyl phosphonate modifications such as methyl-phosphonate-type or ethyl-phosphonate-type, phosphorothioate or phosphoroamidate-modified products, etc. may be
15 used.

[0138]

The term "antisense oligonucleotide(s)" as used herein means, not only those in which the nucleotides corresponding to those constituting a specified region of a DNA or mRNA are entirely
20 complementary, but also those having a mismatch of one or more nucleotides, as long as the DNA or mRNA and the oligonucleotide can selectively and stably hybridize with the nucleotide sequence of SEQ ID NO: 1.

[0139]

25 "Selectively and stably hybridize" means that significant cross hybridization with DNA encoding other proteins does not occur under usual hybridization conditions, preferably under stringent hybridization conditions. Such DNAs are indicated as those having, in the "at least 15 continuous nucleotide" sequence region, a homology
30 of at least 70% or higher, preferably 80% or higher, more preferably 90% or higher, even more preferably 95% or higher nucleotide sequence homology. The algorithm stated herein can be used to determine homology. Such DNA is useful as a probe for detecting or isolating DNA encoding the protein of the invention, or as a primer for
35 amplification as described in Examples below.

[0140]

The antisense oligonucleotide derivative of the present invention acts upon cells producing the protein of the invention by binding to the DNA or mRNA encoding the protein to inhibit its transcription or translation, and to promote the degradation of mRNA, and has an effect of suppressing the function of the protein of the invention by suppressing the expression of the protein.

[0141]

The antisense oligonucleotide derivative of the present invention can be made into an external preparation such as a liniment and a poultice by mixing with a suitable base material, which is inactive against the derivatives.

[0142]

Also, as needed, the derivatives can be formulated into tablets, powders, granules, capsules, liposome capsules, injections, solutions, nose-drops, and freeze-dried agents by adding excipients, isotonic agents, solubilizers, stabilizers, preservatives, pain-killers, etc. These can be prepared using the usual methods.

[0143]

The antisense oligonucleotide derivative is given to the patient by directly applying onto the ailing site, by injecting into a blood vessel, etc. so that it will reach the ailing site. An antisense-mounting material can also be used to increase durability and membrane-permeability. Examples are, liposome, poly-L lysine, lipid, cholesterol, lipofectin, or derivatives of these.

[0144]

The dosage of the antisense oligonucleotide derivative of the present invention can be adjusted suitably according to the patient's condition and used in desired amounts. For example, a dose range of 0.1 to 100 mg/kg, preferably 0.1 to 50 mg/kg can be administered.

[0145]

The antisense oligonucleotide derivative of the present invention is useful in inhibiting the expression of the protein of the invention, and therefore is useful in suppressing the biological activity of the protein of the invention. Also, expression-inhibitors comprising the antisense oligonucleotide derivative of the present invention are useful because of their capability to suppress the biological

activity of the protein of the invention.

[0146]

[Examples]

The present invention shall be described in detail below with
5 reference to examples, but is not be construed as being limited thereto.

[0147]

[Example 1] Two step Blast Search

Probe sequences (256 types) comprising the tggag(t/c)nnntggag(t/c)
(where n is an arbitrary nucleotide) as the oligonucleotide encoding
10 the Trp-Ser-Xaa-Trp-Ser motif were designed. These sequences enable
the detection of almost all known hemopoietin receptors, except for
the EPO receptor, TPO receptor, and the mouse IL6 receptor. Using
each sequence as the query, the GenBank nr database was searched using
the BlastN (Advanced BlastN 2.0.4) program. Default values
15 (Descriptions=100, Alignments=100) were used as parameters for the
search, except for making the expectation value 100.

[0148]

Since approximately 500 clones that completely matched the probe
sequences were obtained as a result of the primary search, among these,
20 a 180-residue nucleotide sequence of human genome-derived clones
(cosmid, BAC, and PAC) containing the probe sequence in approximately
the center was excised. Next, using this 180-residue nucleotide
sequence as the query, the nr database was searched again using the
BlastX (Advanced BlastX 2.0.4) program to search the homology of the
25 amino acid sequence around the probe sequence with known hemopoietin
receptors.

[0149]

Default values were used as parameters for the search, except for
making the expectation value 100. However, when extremely large number
30 of hits were obtained (caused by the Alu sub family that is a high
repetitive sequence), it was often difficult to observe hits for known
hemopoietic receptors. Therefore, to maximize the sensitivity in such
cases, a value of "Expect=1000, Descriptions=500, Alignments=500"
was used.

35 [0150]

As a result of the secondary search by BlastX, 28 clones hit one

or more known hemopoietin receptors (Table 1 to Table 8).

[0151]

[Table 1]

Probe	Xaa	Accession No.	Hit site	Locus	blastx (expect=100)
TGGAGTAATTGGAGC	Asn	AL009181	30892 tggagtaattggagc 30878	lp34.1-1p35	mIL11R(opposite), FOHR
TGGAGCTGATGGAGC	***	Z97987	140006 tggagctgatggagc 139992	lp36.2-36.3	line1, Leu Zip p40,
TGGAGCAGCTGGAGC	Ser	AF023268	39931 tggagcagctggagc 39917	1q21	metaxin
TGGAGCTGCTGGAGC	Cys	AL009051	78023 tggagctctggagc 78037	1q23-24	HP-10, semaphorin F,G
TGGAGCAGCTGGAGT	Thr	Z97200	112905 tggagcagctggagc 112891	1q24	AFP enhancer BP, RAR
TGGAGTGCTGGAGC	Ala	U95626	101031 tggagtgctggagc 101017	3	CFTC, TcR
TGGAGTAGATGGAGT	Arg	Z84495	2547 tggagtagatggagc 2533	3p21.3	trithorax
TGGAGCTGATGGAGT	***	Z74023	5255 tggagctgatggagc 5241	3p21.3	E2ABP, fibronectin, nidgen
TGGAGTTCTGGAGT	Phe	Z68275	7291 tggagttctggagc 7277	4p16.3	mena, NMDAR
TGGAGTGCTGGAGT	Ala	Z54072	21277 tggagtgctggagc 21291	4p16.3	crk, AchR, HER3
TGGAGCTGCTGGAGC	Cys	Z69837	30266 tggagctgctggagc 30252	4p16.3	KIT, FLT3, PDGFRa
TGGAGTTACTGGAGT	Tyr	AC003951	27290 tggagttactggagc 27304	5	collagen
TGGAGCCTGTGGAGT	Leu	AC004502	48334 tggagcctgtggagc 48320	5	ADAMTS-1, properdin, etc
TGGAGTTGATGGAGC	***	L81613	2418 tggagttgatggagc 2404	5	APC, bat2, p53
TGGAGTGATGGAGT	Val	AC002122	43679 tggagtgatggagc 43665	5p15.2	Met tRNA synthase
TGGAGTCCATGGAGT	Pro	AC002380	34646 tggagtcctggagc 34632	5p15.2	N-WASP, enigma
TGGAGCAACTGGAGC	Asn	AC002479	80443 tggagcaactggagc 80457	5p15.2	NEU, glycoprotein C
TGGAGCTGCTGGAGT	Cys	AC004592	125445 tggagctgctggagc 125431	5q31	CD22-B
TGGAGTAGCTGGAGT	Ser	AC002393	3721 tggagtagctggagc 3735	6	glycoprotein
TGGAGTTGCTGGAGT	Cys	AC002326	114578 tggagttgctggagc 114564	6	G3P REGULON
TGGAGTGCTGGAGT	Ala	Z84490	20244 tggagtgctggagc 20230	6	Alu, adrenergic receptor

[0152]
[Table 2]

Probe	Xaa	Accession No.	Hit site	Locus	blastx (expect=100)
TTGGAGTTTCTGGAGC	Phe	AC002112	68699 tggagttcttgagc 68686	6	IgHv, MYD116
TGGAGCGGCTGGAGC	Gly	U89336	35829 tggagcgctggagc 35815	6p21	myosin HC, cep250,
TGGAGCGTCTGGAGC	Val	U53588	3558 tggagcgcttgagc 3572	6p21.3	ring finger, BRCA1
TGGAGTGCATGGAGT	Ala	Z98744	38358 tggagtcattggagt 38344	6p21.3-22.3	Alu, AD7c-NTP
TGGAGTTGCTGGAGT	Cys	AL009031	104325 tggagttgctgagt 104311	6p22.3-24.1	ACC synthase
TGGAGTGTCTGGAGT	Val	AL008729	21325 tggagtgctggagt 21339	6p24	E1A, DUB-2
TGGAGTTGTTGGAGT	Cys	Z98755	69825 tggagttgtggagt 69811	6q16.1-21	dynein
TGGAGCTTCTGGAGC	Phe	Z98172	35554 tggagctcttgagc 35540	6q21	HGXPRT
TGGAGCAGCTGGAGC	Arg	Z97989	79116 tggagcagtgagc 79102	6q21-22	syn fyn, slk, yes, src
TGGAGCTAATGGAGT	***	Z95326	16562 tggagctaatggagt 16576	6q22.1-6q22.33	tyrosinase
TGGAGCTCTTGGAGC	Ser	Z98049	25800 tggagctcttgagc 25786	6q26-q27	collagen, AT3, C1Qb
TGGAGCTCCTGGAGT	Ser	AC003090	22068 tggagctccttgagt 22052	7p15	ICE
TGGAGTATATGGAGC	Ile	AC004744	22740 tggagtatatggagc 22754	7p15-p21	TSH-R, RNABP
TGGAGTAGCTGGAGC	Ser	AC004485	86356 tggagtagctggagc 86370	7p15-p21	Hox 2.4, <u>mLL1Ra(stop*)</u>
TGGAGTCTTTGGAGT	Leu	AC004141	3130 tggagctcttgagt 3144	7p21-p22	polyprotein
TGGAGCAGATGGAGC	Arg	AC004548	62876 tggagcagatggagc 62862	7q11.23-q21.1	NCAM
TGGAGCAACTGGAGT	Asn	AC002456	69500 tggagcaactggagt 69514	7q21	glycoprotein A
TGGAGTAACTGGAGT	Asn	AC000064	9170 tggagtaactggagt 9184	7q21-22	GA3PD
TGGAGTTATTGGAGT	Tyr	AC003085	87341 tggagttattggagt 87355	7q21-22	Nmyc, FGFR
TGGAGTTGTTGGAGT	Cys	AC000119	65235 tggagttgttgagt 65221	7q21-7q22	FVIII, TopoIII
TGGAGTTGTTGGAGT	Cys	AC002458	44435 tggagttgttgagt 44421	7q21-q22	telomerase, NEAT
TGGAGTACATGGAGC	Thr	AC000059	9977 tggagtacatggagc 9963	7q21-7q22	Alu, Notch4

[0153]
[Table 3]

Probe	Xaa	Accession No.	Hit site	Locus	blastx (expect=100)
TGGAGTCAGTGGAGC	Gln	AC000119	103658 tggagtcattggagt 103659	7q21.7q22	reverse transcriptase
TGGAGTATTGGAGT	Ile	AC002384	52216 tggagtattggagt 52202	7q22	pol, GHR(another frame)
TGGAGCAGCTGGAGT	Ser	AC004822	55291 tggagcagctggagt 55277	7q22-q31.1	hemoglobin beta
TGGAGTGTTGGAGT	Val	AC002466	43273 tggagtgttggagt 43287	7q31	ryanodine receptor, mTPO
TGGAGTGGCTGGAGC	Gly	AC002543	112948 tggagtggctggagc 112962	7q31.2	EGF, P-selectin
TGGAGCTGATGGAGC	***	AC000061	79564 tggagctgatggagc 79550	7q31.2	laminin B1, tubulin
TGGAGTTTTGGAGT	Phe	AC000125	13750 tggagttttggagt 13736	7q31.3	p150
TGGAGTTGTTGGAGT	Cys	AC002498	20166 tggagtgttggagt 20162	7q31.3	IL3Rb(anomalous)
TGGAGCGGCTGGAGC	Gly	U66059	158491 tggagcggctggagc 158477	7q35(TcRb)	properdin
TGGAGCATTTGGAGC	Ile	AC003109	4761 tggagcattggagc 4775	7q36	CD2, HOX-2.6
TGGAGTTATTGGAGT	Tyr	AF027390	174448 tggagtattggagt 174434	7q tel	IkB, V2R
TGGAGCATATGGAGT	Ile	AC002052	28882 tggagcattggagt 28896	9p22	myosin VIIA, OSMIR
TGGAGCAACTGGAGT	Asn	AC001643	27345 tggagcaactggagt 27331	9q34	hox1.4, gastrinR
TGGAGCGGATGGAGC	Gly	AC000396	16394 tggagcggatggagc 16380	9q34	vWf, laminin a3
TGGAGTGAGTGGAGT	Glu	U73649	16850 tggagtgaaggagt 16836	11	zinc finger
TGGAGTGGGTGGAGT	Gly	U73649	16850 tggagtgaaggagt 16836	11	zinc finger
TGGAGTGCCCTGGAGT	Ala	U73629	31027 tggagtgcctggagt 31041	11	Alu, gp2b, BCGF-12
TGGAGTCCCTGGAGT	Pro	U73629	36731 tggagtccctggagt 36745	11	ELZRN, hemagglutinin
TGGAGTCCCTGGAGC	Pro	U73643	14550 tggagtccctggagc 14564	11	reverse transcriptase
TGGAGCAACTGGAGC	Asn	AF015116	65621 tggagcaactggagc 65635	11p15.5	Nasopressin R, OSMIR
TGGAGTGCATGGAGT	Ala	AC002350	23543 tggagtgcattggagt 23529	12q24	Alu, IFNaR

[0154]
[Table 4]

Probe	Xaa	Accession No.	Hit site	Locus	blastx (expect=100)
TGGAGTGCATGGAGT	Ala	AC004217	88222 tggagtgcatggagt 88808	12q24.1	Alu, HPK
TTGGAGTTACTGGAGC	Tyr	AC002978	65893 tggagttactggagc 65907	12q24	clathrin LC, EPOR(nonWS)
TGGAGTTGTTGGAGT	Cys	AC000403	91715 tggagttgtggagt 91729	13	VHL, inhibin B
TGGAGCGGTTGGAGC	Gly	X97051	73621 tggagcggttggagc 73607	14q32.33 (IgD)	polycystic kidney
TGGAGTAGGTGGAGC	Arg	AC003024	15596 tggagtaggtggagc 15552	15q26	pksF
TGGAGTTTCTGGAGC	Phe	AC002492	93356 tggagttctggagc 93370	16	pol, PROR
TGGAGTTCATGGAGT	Ser	U91318	102406 tggagttcatggagt 102392	16	ICAM1, MIBP1
TGGAGTGATGGAGT	Val	AC002289	10631 tggagtgtatggagt 10645	16	Alu
TGGAGCCACTGGAGT	His	U91318	162252 tggagccactggagt 162286	16	laminin alpha5
TGGAGTTAATGGAGT	***	AC002519	81768 tggagttaatggagt 81754	16	Rho, Notch
TGGAGCTGCTGGAGT	Cys	U91326	84127 tggagctgctggagt 84113	16p11.2	NIP1-like, IL2Rr(nonWS)
TGGAGTGAATGGAGT	Gln	AC002303	10952 tggagtgaatggagt 10956	16p12	TPOR, OBR, and many
TGGAGCACTTGGAGC	Thr	AC002551	82245 tggagcacttggagc 82259	16p12.1	envelope, androgen R
TGGAGTCCCTGGAGC	Pro	AC002299	162 tggagtccttggagc 148	16p12-p13.1	CYCLIN H, FN
TGGAGCTATTGGAGC	Tyr	AC002299	84540 tggagctattggagc 84526	16p12-p13.1	Alu, RNA editase
TGGAGTCACTGGAGT	His	U95737	16130 tggagtcactggagt 16144	16p13.1	TcRa, HLAA
			16374 tggagtcactggagt 16388		Notch, Pro-rich
			16599 tggagtcactggagt 16613		phosphatase, ORFB
TGGAGTCCCTGGAGC	Pro	U91318	112272 tggagtccttggagc 112286	16p13.1	CD30, collagen, MAP1a
TGGAGCACTTGGAGC	Thr	AC004509	26031 tggagcacttggagc 26045	16p13.3	TcRb
TGGAGCCGTTGGAGC	Arg	AC004496	28217 tggagccgttggagc 28231	16p13.3	mucin, ET1, IL12R(nonWS)

[0155]
[Table 5]

Probe	Xaa	Accession No.	Hit site	Locus	blastx (expect=100)
TGGAGCCGCTGGAGC	Arg	AC004232	34550 tggagcgtggagc 34564	19p13.3	IgLk, AGPR
TTGGAGTACTTGGAGC	Thr	AJ003147	151180 tggagtacttggagc 151166	19p13.3	RanBP2
TGGAGCGTGTGGAGC	Val	X71874	11520 tggagcgttggagc 11534	19q22.1	collagen a5IV
TGGAGCAAATGGAGT	Lys	AC003663	114346 tggagcaaatggagc 114360	17	beta-D-glucosidase
TGGAGTCTCTGGAGC	Leu	AC003957	52898 tggagtctctggagc 52884	17	TTE-1, SEX, Rho,
TGGAGCAGATGGAGC	Arg	AC003971	76277 tggagcagatggagc 76263	18	LIMK-1, TcR
TGGAGTGCATGGAGT	Ala	AD000812	30891 tggagtgcattggagc 30905	19	Alu
TGGAGTGCATGGAGT	Ala	AC002126	86832 tggagtgcattggagc 86846	19	Alu, ADP-NTF
TGGAGTGCCTGGAGT	Cys	AC004660	10008 tggagtgcctggagc 10022	19	Reps1
TGGAGCCCTGGAGT	Pro	AC004490	14389 tggagccctggagc 14403	19	mucin, ataxin-2, N-WASP
TGGAGTGAGTGGAGC	Gln	AC003112	18315 tggagtgaattggagc 18301	19p12(NIR6)	TPOR, PRLR, OBR, etc
TGGAGCAGATGGAGC	Arg	AC004004	39010 tggagcagatggagc 38996	19p12	PRLR, IL12R, GM-
			-----presumably a pseudogene-----		CSERb, IL11R(+stop codon)
TGGAGCACCTGGAGT	Thr	AD000685	39177 tggagcagatggagc 39163		IL3Ra(weak, 22 nonWS)
TGGAGCTGATGGAGC	***	AC002115	21015 tggagcaccctggagc 21001	19p13.1	GM-CSFRb(nonWS+stop)
TGGAGCCAGTGGAGC	Gln	M63796	37164 tggagctgatggagc 37178	19q13.1	Mpc2, Pro rich protein
TGGAGTTACTGGAGT	Tyr	AC004505	7622 tggagccagctggagc 7636	19q13.3	NFCP, titin, Jagged 2
TGGAGTTGATGGAGC	***	Z93016	31711 tggagtacttggagc 31725	20	Gap junction
TGGAGTCAATGGAGT	Gln	135677	31093 tggagtgaattggagc 31079	20q12-13.2	smaphorin F, GHS-R, JAK2
TGGAGTCCCTGGAGT	Ala	AF039907	579 tggagtcaattggagc 565	21(MX1)	GLI, IL13R, IL17R(nonWS)
TGGAGTGTCTGGAGT	Val	AG000937	29892 tggagtgcctggagc 29906	21	IgV, Cyt.Oxidase
			105 tggagtgtctggagc 91	21q	peroxidase

[0156]
[Table 6]

Probe	Xaa	Accession No.	Hit site	Locus	blastx (expect=100)
TGGAGTAAATGGAGT	Lys	AP000034	28803 tggagtaaatggagt 28789	21q11.1	Na/Ca exchanger
TTGGAGTAGGTGGAGT	Arg	AP000039	24900 tggagtagtggagt 24914	21q11.1	RNA polymerase
TGGAGTGAGTGGAGT	Glu	AP000035	21721 tggagtgagtggagt 21707	21q11.1	semaphorin F
TGGAGTGTCTGGAGT	Val	AG000038	26164 tggagtgtctggagt 26150	21q11.1	Glycoprotein
TGGAGTGCCCTGGAGT	Ala	AP000045	7204 tggagtgcctggagt 7218	21q11.1	IgV,
TGGAGCATTTGGAGC	Ile	AP000052	93726 tggagcatttggagc 93740	21q11.1	Ig H, TCF-3, CETP
TGGAGCCTCTGGAGC	Leu	AP000037	17581 tggagcctctggagc 17567	21q11.1	Alu, BCGF
TGGAGTGGGTGGAGT	Gly	AP000016	48480 tggagtgggtggagt 48494	21q22.2	TPO
TGGAGTGAGTGGAGT	Glu	Z97055	151632 tggagtgagtggagt 151618	22	semaphorin H, CD44
TGGAGCTGGTGGAGT	Trp	Z83856	8503 tggagctggtggagt 8489	22	ERF
TGGAGTGGGTGGAGT	Gly	Z95113	69325 tggagtgggtggagt 69311	22q11.2-qter	factor H
TGGAGTGCATGGAGT	Ala	Z93784	36348 tggagtgcattggagt 36362	22q11.2-qter	Alu, NF2
TGGAGCCTCTGGAGT	Leu	AC002308	130741 tggagcctctggagt 130727	22q11.2	collagen a1, Na channel
TGGAGTCCCTGGAGC	Pro	AC000086	40705 tggagtcctggagc 40691	22q11.2	ADH, collagen
TGGAGCATCTGGAGC	Ile	L77569	21088 tggagcatctggagc 21074	22q11.2	Georgelathrin heavy chain 2
TGGAGCAGCTGGAGC	Ser	AC000092	9817 tggagcagctggagc 9803	22q11.2	IgHv, PC binding
TGGAGCAACTGGAGC	Asn	Z95116	64481 tggagcaactggagc 64495	22q12.1	p150, IL4RWSNWSF*
TGGAGCTAGTGGAGC	***	AC003071	114780 tggagctagtggagc 114794	22q12.1-qter	FGFRb
TGGAGCCCTTGGAGC	Pro	Z80902	2675 tggagccctggagc 2661	22q12-qter	collagen a1
TGGAGCTCTTGGAGT	Ser	Z79999	40825 tggagctcttggagt 40839	22q12-qter	collagen a1,

[0157]
[Table 7]

Probe	Xaa	Accession No.	Hit site	Locus	blastx (expect=100)
TGGAGCCATTGGAGT	His	Z81308	12575 tggagccattggagt 12561	22q12-qter	MYF-5, p53, INK4a
TGGAGGAGCTGGAGT	Glu	AL008337	85322 tggagggagctggagt 85336	22q12.3-13.2	GMI-CSFRb, IL3R-EPOR, etc
TTGGAGTGAAGTGGAGT	Glu	U62317	77740 tggagtgagtgagg 77726	22q13	latrophilin-related
TGGAGTGCATGGAGT	Ala	Z98015	31082 tggagtgcatggagt 31068	22q13	Alu, C-CSFR, AD7c-NTP
TGGAGTTGTTGGAGT	Cys	AC002422	19151 tggagttgtggagt 19137	X	cGMP PDase
TGGAGTGTCTGGAGT	Val	Z73418	31830 tggagtgtctggagt 31816	X	WNT-8D, Mi-2
TGGAGTCTTTGGAGT	Leu	Z63843	114972 tggagtctttggagt 114958	X	reverse transcriptase
TGGAGTCTCTGGAGT	Leu	Z99706	7749 tggagtctctggagt 7735	X*	Selenoprotein
TGGAGCAACTGGAGT	Asn	AC002420	70704 tggagcaactggagt 70690	X	homeoprotein, OBR(stop)
TGGAGCATGTGGAGT	Met	Z77239	5702 tggagcatgtggagt 5688	X	TcRb, P14NKR
TGGAGTTCCTGGAGC	Ser	Z83131	4904 tggagtctctggagc 4890	X	VPS41 homolog
TGGAGTGGCTGGAGC	Gly	AC004388	239975 tggagtggctggagc 239989	X	GAP, mLIFR(stop)
TGGAGTGGCTGGAGC	Gly	AC004478	73509 tggagtggctggagc 73495	X	RNAse, mLIFR(stop)
TGGAGTCTATGGAGC	Leu	Z70050	9934 tggagtctatggagc 9948	X	complement C8, C7
TGGAGTCTATGGAGC	Leu	Z73986	40768 tggagtctatggagc 40750	X	complement C8, C7
TGGAGCTGTTGGAGC	Cys	L44140	112657 tggagctgtggagc 112671	X	rab GDI alpha, BDGF
TGGAGCTCATGGAGC	Ser	AC004383	144906 tggagctcatggagc 144892	X	RTase, transposon
TGGAGTAAATGGAGC	Lys	Z69732	31681 tggagtaaatggagc 31695	Xp11	OT-R, acrosin
TGGAGTTCTGGAGC	Ser	Z92545	88703 tggagtctctggagc 88717	Xp11	PMK1
TGGAGCTTCTGGAGC	Phe	AL008709	46089 tggagcttctggagc 46075	Xp11.23-Xp11.4rMHC class 1a, HLA-C	
TGGAGTTTCTGGAGT	Phe	U96409	116332 tggagtttctggagt 116346	Xp22	myosin H
TGGAGTTGCTGGAGT	Cys	AC008105	89544 tggagtgtctggagt 89530	Xp22	IL3R

[0158]
[Table 8]

Probe	Xaa	Accession No.	Hit site	Locus	blastx (expect=100)
TTGGAGTCACTGGAGT	His	AL021706	11982 tggagtcactggagt 11968	Xq21.1-21.33	dopamine receptor
TGGAGCTGGTGGAGT	Trp	AC000113	119188 tggagctgtggagt 119202	Xq23	DNA repair protein, MHC
TGGAGCAAGTGGAGT	Lys	AF007262	98212 tggagcaagtggagt 98226	Xq28	RNA polymerase
TGGAGCTGCTGGAGT	Cys	U82671	35792 tggagctgtggagt 35806	Xq28	XTCF-3c
TGGAGTCAGTGGAGC	Gln	AF011889	144465 tggagtcagtggagc 144461	Xq28	GHRHR, Werner Synd.
TGGAGCTAATGGAGC	***	AF030876	107409 tggagctaaggagc 107395	Xq28	gp41, clk3
		AF031078	94438 tggagtaagtggagc 94424	Xq28	gp41
		Z47046	5688 tggagtaagtggagc 5692	Xq28	gp41
TGGAGTTTCTGGAGT	Phe	AC002531	106698 tggagtttctggagt 106712	Y	Alu, bpk
TGGAGCAGTTGGAGC	Ser	AC004474	124745 tggagcagtggagc 124731	Y	EGFR, Smad6
TGGAGTTTGTGGAGT	Leu	U26425	12899 tggagtttgtggagt 12913	PLCb2	PRLR(opposite)
TGGAGCAACTGGAGT	Asn	U96726	61672 tggagcaactggagt 61658	mouse DNA	envelope mIL11R(opposite)
TGGAGTCCCTGGAGC	Pro	<u>U35321</u>	22244 tggagtcctggagc 22230	MHC class II	CFTC, <u>IL6R</u>
TGGAGCAGATGGAGC	Arg	AC002482	14276 tggagcagatggagc 14290	RG208O03	I-309, TcR, IL9R(nonWS)
TGGAGCTCTTGGAGC	Ser	U34879	24914 tggagctcttggagc 24928	EDH17B2	Large tegument protein
TGGAGCCTTTGGAGC	Leu	Z15025	6359 tggagcctttggagc 6373	Bat2	commonB(opposit.nonWS)
GM-CSFRb(opposite.stop)					bat2,mucin,

Redundant clones are shadowed. White and underlined letters indicate hits and pseudo-hits, respectively.

[0159]

Four clones out of these 28 clones (AC002303, AC003112, AL008637, and AC004004) hit several known hemopoietin receptors, however, AC004004 was excluded as it has a stop codon downstream three amino acids of the Trp-Ser-Xaa-Trp-Ser motif. Among the three remaining clones, AL008637 was thought to be a known receptor, GM-CSF receptor β . AC002303 is the BAC clone CIT987-SKA-670B5 derived from the 16p12 region of human chromosome no. 16 registered by TIGR group on June 19, 1997 and comprises the full-length of 131530 base pairs (Lamerdin, J.E., et al., GenBank Report on AC003112, 1997).

[0160]

As shown in Fig. 1, a BlastX search (query: 180 nucleotides of 40861-41040 including tggagtgaatggagt (40952-40966), the only probe sequence within the AC002303) revealed that numerous hemopoietin receptors starting with the TPO receptor and leptin receptor show an evident homology, however, there were no known, database-registered hemopoietin receptors that completely matched the query sequence. Also, a BlastX scanning was done under the above conditions, by excising a sequential 180-residue nucleotide sequence in both the 5' and 3' directions, centering on the 180-residue nucleotide sequence mentioned above, and when this was used as a query, two sequences having a homology to known hemopoietin receptors were found in the regions 39181-39360 and 42301-42480, and were thought to be other exons of the same gene (Fig. 2).

[0161]

A Pro-rich motif PAPPF was conserved in the 39181-39360 site, and a Box 1 motif in the 42301-42480 site. The 3' side exon adjacent to the exon containing the Trp-Ser-Xaa-Trp-Ser motif has a transmembrane domain, and this domain has a low homology with other hemopoietin receptors, and was not detected by the BlastX scan. These results suggested the possibility of a novel hemopoietin receptor gene existing in the above-described BAC clone CIT987-SKA-670B5.

[0162]

[Example 2] Search for NR8 expressing tissues using RT-PCR

Pseudogenes have been reported to exist in several hemopoietin receptors (Kermouni, A. et al., Genomics, 1995, 29 (2) 371-382; Fukunaga,

R. and Nagata, S., *Eur. J. Biochem.*, 1994, 220, 881-891). To verify that NR8 is not a pseudogene, and with the objective of identifying NR8 expressing tissues, transcripts of the NR8 gene were searched by RT-PCR method.

5 [0163]

In the AC002303 sequence of the above-described BAC clone, several exon regions widely conserved at the amino acid translation level in known cytokine receptors were surmised, and on the sequence of the surmised exon region, the following primers were synthesized.

10 (See Fig. 5 for the location of each primer.)

NR8-SN1; 5'- CCG GCT CCC CCT TTC AAC GTG ACT GTG ACC -3' (SEQ ID NO: 9)

NR8-SN2; 5'- GGC AAG CTT CAG TAT GAG CTG CAG TAC AGG -3' (SEQ ID NO: 10)

15 NR8-AS1; 5'- ACC CTC TGA CTG GGT CTG AAA GAT GAC CGG -3' (SEQ ID NO: 11)

NR8-AS2; 5'- CAT GGG CCC TGC CCG CAC CTG CAG CTC ATA -3' (SEQ ID NO: 12)

Using the Human Fetal Multiple Tissue cDNA Panel (Clontech #K1425-1) as the template, RT-PCR was attempted using combinations of the above primers. Advantage cDNA Polymerase Mix (Clontech #8417-1) was used for the PCR, which was conducted under the conditions below using the Perkin Elmer Gene Amp PCR System 2400 Thermalcycler.

[0164]

25 Namely, the PCR conditions were, 94°C for 4 min, 5 cycles of "94°C for 20 sec, 72°C for 3 min," 5 cycles of "94°C for 20 sec, 70°C for 3 min," 28 cycles of "94°C for 20 sec, 68°C for 3 min," 72°C for 4 min, and completed at 4°C.

[0165]

30 From the primer locations shown in Fig. 5, amplifications of bands sized 330 bp, 258 bp, 234 bp, and 162 bp can be expected from the combinations of SN1/AS1, SN1/AS2, SN2/AS1, and SN2/AS2. When evaluated using human fetal liver, brain, and skeletal muscle cDNA as the template, clear bands having the anticipated sizes were obtained
35 in the fetal liver only with the respective primer combinations (Fig. 3).

[0166]

An amplification was not seen at all for fetal brain cDNA, and a band of about 650 bp and a broad band of 400 to 500 bp were observed for fetal skeletal muscle cDNA. However, since the band sizes for
5 skeletal muscle cDNA remained constant even when different combinations of primers were used, it is thought that these bands were non-specific amplifications due to some reason.

[0167]

The obtained PCR product was subcloned to pGEM-T Easy vector (Promega
10 #A1360), and the nucleotide sequence was determined. The recombination of PCR products to the pGEM-T Easy vector was done by T4 DNA Ligase (Promega #A1360) reacted at 4°C for 12 hr. The genetic recombinant between the PCR product and pGEM-T Easy vector was obtained by transforming *E. coli* strain DH5α (Toyobo #DNA-903).

15 [0168]

For the selection of the genetic recombinant, Insert Check Ready (Toyobo #PIK-101) was used. The dRhodamine Terminator Cycle Sequencing Kit (ABI/Perkin Elmer #4303141) was used for determining the nucleotide sequence, and analysis was done using the ABI PRISM
20 377 DNA Sequencer. As a result of determining the nucleotide sequences of all inserts of the 10 independent clones of genetic recombinants, all clones were found to comprise a single nucleotide sequence. These obtained sequences were verified to be partial nucleotide sequences of NR8.

25 [0169]

[Example 3] Full-length cDNA cloning by the 5' and 3'-RACE methods

Using the thus-obtained fetal liver-derived cDNA, 5' and 3'-RACE methods were conducted to obtain full-length cDNA (Fig. 4).

[0170]

30 3-1) 5'-RACE method

5'-RACE PCR was performed using the above-mentioned NR8-AS1 primer for primary PCR, and NR8-AS2 primer for secondary PCR. Human Fetal Liver Marathon-Ready cDNA Library (Clontech #7403-1) was used as the template and Advantage cDNA Polymerase Mix for the PCR experiment.
35 As a result of PCR under the following conditions using the Perkin Elmer Gene Amp PCR System 2400 Thermalcycler, two types of PCR products

were obtained, which have different sizes through selective splicing.

[0171]

Primary PCR conditions were 94°C for 4 min, 5 cycles of "94°C for 20 sec, 72°C for 4 min," 5 cycles of "94°C for 20 sec, 70°C for 4 min,"
5 28 cycles of "94°C for 20 sec, 68°C for 4 min," 72°C for 4 min, and completed at 4°C.

[0172]

Secondary PCR conditions were 94°C for 4 min, 5 cycles of "94°C for 20 sec, 70°C for 3 min 30 sec," 28 cycles of "94°C for 20 sec, 68°C
10 for 3 min 30 sec," 72°C for 4 min, and completed at 4°C.

[0173]

Both types of PCR products obtained were subcloned to pGEM-T Easy vector as mentioned earlier, and the nucleotide sequences of all inserts were determined for the 16 independent clones of genetic transformants.
15 As before, the dRhodamine Terminator Cycle Sequencing Kit was used for determining the nucleotide sequence, and analysis was done using the ABI PRISM 377 DNA Sequencer. As a result, the clones can be divided into two groups, one having 14 clones, and the other having 2 clones, by the length of the base pairs and the differences in sequence (though
20 described later, the differences lie in the products due to selective splicing, and the group of 14 independent clones comprises the sequence corresponding to exon 5 in the genomic sequence, and the remaining group of two independent clones does not have this sequence).

[0174]

25 3-2) 3'-RACE method

3'-RACE PCR was performed using the above-mentioned NR8-SN1 primer for primary PCR, and NR8-SN2 primer for secondary PCR. Human Fetal Liver Marathon-Ready cDNA Library was used as the template similar to 5'-RACE PCR, and Advantage cDNA Polymerase Mix for the PCR experiment.
30 As a result of conducting PCR under the conditions shown in 3-1), a single band PCR product was obtained.

[0175]

The obtained PCR product was subcloned to pGEM-T Easy vector as above, and the nucleotide sequences of all inserts of the 12 independent
35 clones of genetic recombinants were determined. As before, the dRhodamine Terminator Cycle Sequencing Kit was used for determining

the nucleotide sequence, and the sequences determined were analyzed using the ABI PRISM 377 DNA Sequencer. As a result, all 12 independent clones showed a single nucleotide sequence.

[0176]

5 As a result of analyzing the nucleotide sequence of the fragments (approximately 1.1 kb and 1.2 kb) amplified by 5'-RACE and 3'-RACE, respectively, it was conceived that the approximately 260 bp of each fragment overlap and extend to the 5' side and 3' side, and contain almost the full-length of NR8 mRNA. These were joined to make a
10 full-length cDNA (NR8 α) (Fig. 5 and Fig. 6). The plasmid containing the NR8 α cDNA (SEQ ID NO: 2) was named pGEM-NR8 α , and *E. coli* containing the plasmid has been internationally deposited at the National Institute of BioScience and Human-Technology, Agency of Industrial Science and Technology (1-3, Higashi 1-chome, Tsukuba-shi, Ibaraki-ken,
15 Japan 305-8566) under the accession number FERM BP-6543 since October 9, 1998 according to the Budapest Treaty.

[0177]

As shown in Fig. 5 and Fig. 6, in the ORF of NR8 α cDNA, the Met starting from nucleotide no. 441 is thought to be the start codon
20 due to the presence of an inframe stop codon 39 bp upstream, and completes with two stop codons starting from nucleotide no. 1524. It has the features of, from the N terminus in order, a typical secretion signal sequence, a domain thought to be the ligand binding site containing a Cys residue conserved in other hemopoietic receptor members, a
25 Pro-rich motif, Trp-Ser-Xaa-Trp-Ser motif, a transmembrane domain, a Box 1 motif thought to be involved in signal transduction, and such features of hemopoietin receptors. From the above results, the NR8 gene was thought to encode a novel hemopoietin receptor.

[0178]

30 Analysis of fragments amplified by the RACE method suggested the presence of a splice variant. As a result of nucleotide sequence analysis, this variant was revealed to be lacking approximately 150 bp including the above-described Pro-rich motif of NR8 α . Moreover, as a result of comparing AC002303 sequence with NR8 α , and carrying
35 out analogy of exons/introns (Table 9), the above-described variant was thought to be deficient of the 5th exon due to selective splicing.

[0179]
[Table 9]

Exon	# in AC002303	# in NR8	Characteristics
1	<1	: 1-424	inframe stop codon
2	26334-26398	: 425-489	start codon, signal peptide
3	30625-30727	: 490-592	conserved Cys residue
4	33766-33965	: 593-792	conserved Cys residue, N-glycosylation site
5	39240-39394	: 793-947	Pro-rich motif (PAPPF), N-glycosylation site
6	40820-40997	: 948-1125	gtWSEWSdp motif
7	41455-41554	: 1126-1225	transmembrane domain
8	42285-42366	: 1226-1307	Box1 (IWA VPSP)
9a	44812-44909	: 1308-1405*	connects to exon 10, Box2-like sequence (PSTLEVYSCH), nontypical exon/intron boundary
9b	44812-45922<	: 1308-2465**	double stop codons, Box2-like sequence (PSTLEVYSCH, PAELVESDG), polyA
10	45441-45922<	: 1406-1934*	double stop codons, polyA
NR8 α^* : exons 1+2+3+4+5+6+7+8+9a+10			
NR8 β : exons 1+2+3+4+6+7+8+9a+10			
(two alternative reading frames for soluble-type and transmembrane(-signal)-type)			
NR8 γ^{**} : exons 1+2+3+4+5+6+7+8+9b			

[0180]

This variant (NR8 β) can encode a soluble receptor in the truncated form by the joining of the 6th exon directly to the 4th exon and causing a frame shift. The boundary between the exons and the introns takes a consensus sequence in most cases, but the boundary between the 9th exon (Exon 9a) and the 9th intron is the only boundary that takes a different sequence from the consensus sequence (nag/gtgagt, etc.), being acc/acggag. The plasmid comprising NR8 β cDNA (SEQ ID NO: 2) was named pGEM-NR8 β , and *E.coli* comprising the plasmid has been internationally deposited at the National Institute of BioScience and Human-Technology, Agency of Industrial Science and Technology (1-3, Higashi 1-chome, Tsukuba-shi, Ibaraki-ken, Japan) under the accession number FERM BP-6544 since October 9, 1998 according to the Budapest Treaty.

[0181]

[Example 4] Northern blotting

In order to analyze the distribution and mode of NR8 gene expression in each human organ and human cancer cell lines, Northern blot analysis was done using the cDNA encoding the full-length NR8 α protein prepared based on all the cDNA fragments obtained in Example 3 as a probe. The probe was prepared using Mega Prime Kit (Amersham, cat#RPN1607) and radiolabeled with [α -³²P] dCTP (Amersham, cat#AA0005).

[0182]

As Northern blots, Human Multiple Tissue Northern (MTN) Blot (Clontech #7760-1), Human MTN Blot IV (Clontech #7766-1), and Human Cancer Cell Line MTN Blot (Clontech #7757-1) were used. Express Hyb Hybridization Solution (Clontech #8015-2) was used for hybridization.

[0183]

Hybridization conditions were: a prehybridization at 68°C for 30 min, followed by hybridization at 68°C for 14 hr. After washing under the following conditions, the blots were exposed to Imaging Plate (FUJI#BAS-III), and the gene expression of NR8 mRNA was detected by the Image Analyzer (FUJIX, BAS-2000 II). Washing conditions were: (1) 1x SSC/0.1% SDS, at room temperature for 5 min; (2) 1x SSC/0.1% SDS, at 50°C 30 min; and (3) 0.1x SSC/0.1% SDS, at 50°C 30 min.

[0184]

Fig. 12 shows the results of Northern blot analysis of NR8 expression

in each organ. A total of three different-sized mRNA, one 5kb-sized and two 3 to 4kb sized, were detected in human adult lung, spleen, thymus, skeletal muscle, pancreas, small intestines, peripheral leucocytes, and uterus. A similar examination of various cell lines including hemopoietic cell lines showed similar sized bands in two cell lines, the promyeloid leukemic cell line HL60 and Burkett's lymphoma-derived Raji.

[0185]

[Example 5] Plaque screening

Northern blot analysis of NR8 gene expression detected at least three types of specific mRNA bands with different sizes in each human organ and in each human cancer cell line for which NR8 gene expression was seen. However, the inventors had succeeded in isolating only two types of selective splicing variants, namely NR8 α and NR8 β genes, in the above-described Examples. Therefore, the inventors performed plaque screening with the objective of isolating the gene of the third selective splicing variant. Human Lymph Node (Clontech, cat#HL5000a) that showed a strong NR8 gene expression in the above-mentioned Northern analysis results, was used as the cDNA library. The probe used was NR8 α cDNA fragment, which was radio-labeled by [α -³²P] dCTP (Amersham, cat#AA0005) using the Mega Prime Kit (Amersham, cat#RPN1607). Approximately 7.2×10^5 plaques of Human Lymph Node cDNA Library were blotted onto a Hybond N (+) (Amersham, cat#RPN303B) charged nylon membrane to conduct primary screening. Rapid Hybridization Buffer (Amersham, cat#RPN1636) was used for the hybridization. Hybridization conditions were: a prehybridization at 65°C for 1 hr, followed by hybridization at 65°C for 14 hr. After washing under the conditions, (1) 1x SSC/0.1% SDS, at room temperature for 15 min; (2) 1x SSC/0.1% SDS, at 58°C 30 min; and (3) 0.1x SSC/0.1% SDS, at 58°C 30 min, the membrane was exposed to an X-ray film (Kodak, cat#165-1512) to detect NR8 positive plaques.

[0186]

As a result, positive or pseudo-positive 16 independent clones were obtained. A similar secondary screening was done for the 16 clones obtained from the primary screening to successfully isolate plaques of NR8 positive 15 independent clones. The inserts of these 15 clones

were amplified by PCR through a pair of primers located in both ends of the λ gt10 vector cloning site. Advantage cDNA polymerase Mix (Clontech #8417-1) was used for the PCR reaction conducted using the Perkin Elmer Gene Amp PCR System 2400 Thermalcycler, under the following experiment conditions. Namely, 94°C for 4 min, 5 cycles of "94°C for 20 sec, 70°C for 4 min," 30 cycles of "94°C for 20 sec, 68°C for 4 min," 72°C for 4 min, and completed at 4°C.

[0187]

Similar to above, the obtained PCR products were subcloned to pGEM-T Easy vector, and the nucleotide sequence of the inserts were determined using the BigDye Terminator Cycle Sequencing SF Ready Reaction Kit (ABI/Perkin Elmer #4303150), and analyzed by the ABI PRISM 377 DNA Sequencer. As a result, among the 15 clones obtained, at least two clones showed an insertion of 177 amino acids flanking the NR8 α C terminus, and since this portion derives from the 9th intron of the NR8 gene and is removed by splicing in NR8 α , this 3rd selective splicing variant was named NR8 γ . The plasmid containing the NR8 γ cDNA (SEQ ID NO: 2) was named pGEM-NR8 γ , and *E. coli* containing the plasmid has been internationally deposited at the National Institute of BioScience and Human-Technology, Agency of Industrial Science and Technology (1-3, Higashi 1-chome, Tsukuba-shi, Ibaraki-ken, Japan) under the accession number FERM BP-6545 since October 9, 1998 according to the Budapest Treaty.

[0188]

[Example 6] Ligand screening

6-1) Construction of NR8 chimeric receptor

A screening system was constructed for searching a ligand that can specifically bind to NR8, namely, a novel hemopoietin. First, the cDNA sequence encoding the extracellular region of NR8 α (the amino acid sequence of SEQ ID NO: 1; from the 1st Met to the 228th Glu) was amplified by PCR, and this DNA fragment was bound to DNA fragments encoding the transmembrane region and the intracellular region of a known hemopoietin receptor to prepare a fusion sequence encoding a chimeric receptor. As described above, there were several candidates for the partner, the known hemopoietin receptor, and among them, the human TPO receptor (Human MPL-P) was selected. Namely, after

amplifying the DNA sequence encoding the intracellular region that includes the transmembrane region of the human TPO receptor by PCR, this sequence was bound to the cDNA sequence encoding the extracellular region of NR8 α in frame, and inserted into a plasmid vector expressible in mammalian cells. The expression vector constructed was named pEF-NR8/TPO-R. A schematic diagram of the structure of the constructed NR8/TPO-R chimeric receptor is shown in Fig. 14, and the nucleotide sequence of the chimeric receptor and the expressible amino acid sequence encoded by it are shown in SEQ ID NOs: 13 and 14, respectively. Together with an expression vector pSV2bsr (Kaken Pharmaceutical Co., Ltd.) containing Blastcidin S resistant gene, the NR8/TPO-R chimeric receptor-expressing vector was introduced into the growth factor-dependent cell line Ba/F3, and forced expression was done. Gene-introduced cells were selected by culturing with 8 μ g/ml of Blastcidin S hydrochloride (Kaken Pharmaceutical Co., Ltd.) and IL-3. By transferring the obtained chimeric receptor-introduced cells to an IL-3-free medium, adding a material expected to contain a target ligand, and culturing, it is possible to conduct a screening that uses the fact that survival/proliferation will be possible only when a ligand that specifically binds to NR8 is present.

[0189]

6-2) Preparation of NR8/IgG1-Fc soluble fusion protein

NR8/IgG1-Fc soluble fusion protein was prepared to be used for searching cell membrane-bound ligands, or the detection of soluble ligands through BIAcore (Pharmacia) and West-western blotting. A fusion sequence encoding the soluble fusion protein was prepared by binding a DNA fragment encoding the extracellular region of NR8 α (amino acid sequence; from the 1st Met to the 228th Glu) prepared in 5-1) with the DNA fragment encoding the Fc region of human immunoglobulin IgG1 in frame. A schematic diagram of the structure of the soluble fusion protein encoding the NR8/IgG1-Fc is shown in Fig. 14, and the nucleotide sequence and the expressible amino acid sequence encoded by it in SEQ ID NOs: 15 and 16, respectively. This fusion gene fragment was inserted into a plasmid vector expressible in mammalian cells, and the constructed expression vector was named pEF-NR8/IgG1-Fc. After forced expression of this pEF-NR8/IgG1-Fc in mammalian cells, and

selection of stable gene-introduced cells, the recombinant protein secreted into the culture supernatant can be purified by immunoprecipitation using anti-human IgG1-Fc antibody, or by affinity columns, etc.

5 [0190]

6-3) Construction of an expression system of NR8 β and purification of recombinant NR8 β protein

The recombinant NR8 β protein was prepared to be used for searching cell membrane-bound ligands, or the detection of soluble ligands using BIAcore (Pharmacia) or West-western-blotting. Using the amino acid coding sequence of NR8 β cDNA, the stop codon was replaced by point mutation to a nucleotide sequence encoding an arbitrary amino acid residue, and then, was bound to the nucleotide sequence encoding the FLAG peptide in frame. This bound fragment was inserted into a plasmid vector expressible within mammalian cells, and the constructed expression vector was named pEF-BOS/NR8 β FLAG. Fig. 14 shows a schematic diagram of the structure of the insert NR8 β FLAG within the constructed expression vector. Moreover, the nucleotide sequence of NR8 β FLAG and the expressible amino acid sequence encoded by it are shown in SEQ ID NOs: 17 and 18, respectively. After forced-expression of this pEF-BOS/NR8 β FLAG in mammalian cells and selection of stable gene-introduced cells, the recombinant protein secreted into the culture supernatant can be immunoprecipitated using anti-FLAG peptide antibody, or may be purified by affinity columns, etc.

25 [0191]

[Effects of the Invention]

The present invention provides a novel hemopoietin receptor protein, and the encoding DNA. The present invention also provides, a vector into which the DNA has been inserted, a transformant harboring the DNA, and a method of producing a recombinant protein using the transformant. It also provides a method of screening a compound or a natural ligand that binds to the protein. The protein of the invention is thought to be related to hemopoiesis, and therefore, is useful in experiments for analyzing hemopoietic functions. The protein could also be applied in the diagnosis and treatment of hemopoiesis-associated diseases.

[Sequence Listing]

SEQUENCE LISTING

5 <110> CHUGAI RESEARCH INSTITUTE FOR MOLECULAR MEDICINE, INC.

<120> NOVEL HEMOPOIETIN RECEPTOR PROTEIN

<130> C2-004

10

<150> JP 10-214720

<151> 1998-6-24

<160> 18

15

<170> PatentIn version 2.0

<210> 1

<211> 361

20

<212> PRT

<213> Homo sapiens

<400> 1

25

Met Pro Arg Gly Trp Ala Ala Pro Leu Leu Leu
1 5 10

Leu Leu Leu Gln Gly Gly Trp Gly Cys Pro Asp Leu Val Cys Tyr Thr
15 20 25

30

Asp Tyr Leu Gln Thr Val Ile Cys Ile Leu Glu Met Trp Asn Leu His
30 35 40

Pro Ser Thr Leu Thr Leu Thr Trp Gln Asp Gln Tyr Glu Glu Leu Lys
45 50 55

35

Asp Glu Ala Thr Ser Cys Ser Leu His Arg Ser Ala His Asn Ala Thr

	60		65		70		75									
	His	Ala	Thr	Tyr	Thr	Cys	His	Met	Asp	Val	Phe	His	Phe	Met	Ala	Asp
					80					85					90	
5	Asp	Ile	Phe	Ser	Val	Asn	Ile	Thr	Asp	Gln	Ser	Gly	Asn	Tyr	Ser	Gln
				95					100					105		
	Glu	Cys	Gly	Ser	Phe	Leu	Leu	Ala	Glu	Ser	Ile	Lys	Pro	Ala	Pro	Pro
10			110					115					120			
	Phe	Asn	Val	Thr	Val	Thr	Phe	Ser	Gly	Gln	Tyr	Asn	Ile	Ser	Trp	Arg
		125					130					135				
15	Ser	Asp	Tyr	Glu	Asp	Pro	Ala	Phe	Tyr	Met	Leu	Lys	Gly	Lys	Leu	Gln
	140					145					150				155	
	Tyr	Glu	Leu	Gln	Tyr	Arg	Asn	Arg	Gly	Asp	Pro	Trp	Ala	Val	Ser	Pro
				160					165					170		
20	Arg	Arg	Lys	Leu	Ile	Ser	Val	Asp	Ser	Arg	Ser	Val	Ser	Leu	Leu	Pro
			175					180					185			
	Leu	Glu	Phe	Arg	Lys	Asp	Ser	Ser	Tyr	Glu	Leu	Gln	Val	Arg	Ala	Gly
25		190					195					200				
	Pro	Met	Pro	Gly	Ser	Ser	Tyr	Gln	Gly	Thr	Trp	Ser	Glu	Trp	Ser	Asp
		205					210					215				
30	Pro	Val	Ile	Phe	Gln	Thr	Gln	Ser	Glu	Glu	Leu	Lys	Glu	Gly	Trp	Asn
	220					225					230				235	
	Pro	His	Leu	Leu	Leu	Leu	Leu	Leu	Val	Ile	Val	Phe	Ile	Pro	Ala	
				240					245					250		
35	Phe	Trp	Ser	Leu	Lys	Thr	His	Pro	Leu	Trp	Arg	Leu	Trp	Lys	Lys	Ile

	255		260		265												
	Trp	Ala	Val	Pro	Ser	Pro	Glu	Arg	Phe	Phe	Met	Pro	Leu	Tyr	Lys	Gly	
	270						275						280				
5	Cys	Ser	Gly	Asp	Phe	Lys	Lys	Trp	Val	Gly	Ala	Pro	Phe	Thr	Gly	Ser	
	285						290						295				
	Ser	Leu	Glu	Leu	Gly	Pro	Trp	Ser	Pro	Glu	Val	Pro	Ser	Thr	Leu	Glu	
10	300					305					310					315	
	Val	Tyr	Ser	Cys	His	Pro	Pro	Ser	Ser	Pro	Val	Glu	Cys	Asp	Phe	Thr	
					320					325					330		
15	Ser	Pro	Gly	Asp	Glu	Gly	Pro	Pro	Arg	Ser	Tyr	Leu	Arg	Gln	Trp	Val	
				335					340					345			
	Val	Ile	Pro	Pro	Pro	Leu	Ser	Ser	Pro	Gly	Pro	Gln	Ala	Ser			
	350						355					360					
20	<210> 2																
	<211> 1884																
	<212> DNA																
	<213> Homo sapiens																
25	<220>																
	<221> CDS																
	<222> (441).. (1523)																
30	<400> 2																
	ggcagccagc	ggcctcagac	agacccactg	gcgtctctct	gctgagtgac	cgtaagctcg											60
	gcgtctggcc	ctctgcctgc	ctctccctga	gtgtggctga	cagccacgca	gctgtgtctg											120
35	tctgtctgcg	gcccgatgc	ccctgctgcg	gccgcctggt	accttccttg	ccgtctcttt											180

	cctctgtctg ctgctctgtg ggacacctgc ctggaggccc agctgcccgt catcagagtg	240
	acaggtctta tgacagcctg attggtgact cgggctgggt gtggattctc accccaggcc	300
5	tctgcctgct ttctcagacc ctcatctgtc acccccacgc tgaaccacgc tgccaccccc	360
	agaagcccat cagactgccc ccagcacacg gaatggattt ctgagaaaga agccgaaaca	420
10	gaaggcccgt gggagtcagc atg ccg cgt ggc tgg gcc gcc ccc ttg ctc ctg	473
	Met Pro Arg Gly Trp Ala Ala Pro Leu Leu Leu	
	1 5 10	
	ctg ctg ctc cag gga ggc tgg ggc tgc ccc gac ctc gtc tgc tac acc	521
15	Leu Leu Leu Gln Gly Gly Trp Gly Cys Pro Asp Leu Val Cys Tyr Thr	
	15 20 25	
	gat tac ctc cag acg gtc atc tgc atc ctg gaa atg tgg aac ctc cac	569
	Asp Tyr Leu Gln Thr Val Ile Cys Ile Leu Glu Met Trp Asn Leu His	
	30 35 40	
20	ccc agc acg ctc acc ctt acc tgg caa gac cag tat gaa gag ctg aag	617
	Pro Ser Thr Leu Thr Leu Thr Trp Gln Asp Gln Tyr Glu Glu Leu Lys	
	45 50 55	
25	gac gag gcc acc tcc tgc agc ctc cac agg tcg gcc cac aat gcc acg	665
	Asp Glu Ala Thr Ser Cys Ser Leu His Arg Ser Ala His Asn Ala Thr	
	60 65 70 75	
	cat gcc acc tac acc tgc cac atg gat gta ttc cac ttc atg gcc gac	713
30	His Ala Thr Tyr Thr Cys His Met Asp Val Phe His Phe Met Ala Asp	
	80 85 90	
	gac att ttc agt gtc aac atc aca gac cag tct ggc aac tac tcc cag	761
35	Asp Ile Phe Ser Val Asn Ile Thr Asp Gln Ser Gly Asn Tyr Ser Gln	
	95 100 105	

	gag tgt ggc agc ttt ctc ctg gct gag agc atc aag ccg gct ccc cct	809
	Glu Cys Gly Ser Phe Leu Leu Ala Glu Ser Ile Lys Pro Ala Pro Pro	
	110 115 120	
5	ttc aac gtg act gtg acc ttc tca gga cag tat aat atc tcc tgg cgc	857
	Phe Asn Val Thr Val Thr Phe Ser Gly Gln Tyr Asn Ile Ser Trp Arg	
	125 130 135	
	tca gat tac gaa gac cct gcc ttc tac atg ctg aag ggc aag ctt cag	905
10	Ser Asp Tyr Glu Asp Pro Ala Phe Tyr Met Leu Lys Gly Lys Leu Gln	
	140 145 150 155	
	tat gag ctg cag tac agg aac cgg gga gac ccc tgg gct gtg agt ccg	953
	Tyr Glu Leu Gln Tyr Arg Asn Arg Gly Asp Pro Trp Ala Val Ser Pro	
15	160 165 170	
	agg aga aag ctg atc tca gtg gac tca aga agt gtc tcc ctc ctc ccc	1001
	Arg Arg Lys Leu Ile Ser Val Asp Ser Arg Ser Val Ser Leu Leu Pro	
	175 180 185	
20		
	ctg gag ttc cgc aaa gac tcg agc tat gag ctg cag gtg cgg gca ggg	1049
	Leu Glu Phe Arg Lys Asp Ser Ser Tyr Glu Leu Gln Val Arg Ala Gly	
	190 195 200	
	ccc atg cct ggc tcc tcc tac cag ggg acc tgg agt gaa tgg agt gac	1097
25	Pro Met Pro Gly Ser Ser Tyr Gln Gly Thr Trp Ser Glu Trp Ser Asp	
	205 210 215	
	ccg gtc atc ttt cag acc cag tca gag gag tta aag gaa ggc tgg aac	1145
30	Pro Val Ile Phe Gln Thr Gln Ser Glu Glu Leu Lys Glu Gly Trp Asn	
	220 225 230 235	
	cct cac ctg ctg ctt ctc ctc ctg ctt gtc ata gtc ttc att cct gcc	1193
	Pro His Leu Leu Leu Leu Leu Leu Leu Val Ile Val Phe Ile Pro Ala	
35	240 245 250	

	ttc tgg agc ctg aag acc cat cca ttg tgg agg cta tgg aag aag ata	1241
	Phe Trp Ser Leu Lys Thr His Pro Leu Trp Arg Leu Trp Lys Lys Ile	
	255 260 265	
5	tgg gcc gtc ccc agc cct gag cgg ttc ttc atg ccc ctg tac aag ggc	1289
	Trp Ala Val Pro Ser Pro Glu Arg Phe Phe Met Pro Leu Tyr Lys Gly	
	270 275 280	
10	tgc agc gga gac ttc aag aaa tgg gtg ggt gca ccc ttc act ggc tcc	1337
	Cys Ser Gly Asp Phe Lys Lys Trp Val Gly Ala Pro Phe Thr Gly Ser	
	285 290 295	
15	agc ctg gag ctg gga ccc tgg agc cca gag gtg ccc tcc acc ctg gag	1385
	Ser Leu Glu Leu Gly Pro Trp Ser Pro Glu Val Pro Ser Thr Leu Glu	
	300 305 310 315	
20	gtg tac agc tgc cac cca ccc agc agc cct gtg gag tgt gac ttc acc	1433
	Val Tyr Ser Cys His Pro Pro Ser Ser Pro Val Glu Cys Asp Phe Thr	
	320 325 330	
25	agc ccc ggg gac gaa gga ccc ccc cgg agc tac ctc cgc cag tgg gtg	1481
	Ser Pro Gly Asp Glu Gly Pro Pro Arg Ser Tyr Leu Arg Gln Trp Val	
	335 340 345	
30	gtc att cct ccg cca ctt tcg agc cct gga ccc cag gcc agc taa	1526
	Val Ile Pro Pro Pro Leu Ser Ser Pro Gly Pro Gln Ala Ser	
	350 355 360	
35	tgaggctgac tggatgtcca gagctggcca ggccactggg ccctgagcca gagacaaggt	1586
	cacctgggct gtgatgtgaa gacacctgca gcctttggtc tcctggatgg gcctttgagc	1646
	ctgatgttta cagtgtctgt gtgtgtgtgc atatgtgtgt gtgtgcatat gcatgtgtgt	1706
	gtgtgtgtgt gtcttaggtg cgcagtgga tgtccacgtg tgtgtgattg cacgtgcctg	1766

tgggcctggg ataatgccca tgg tactcca tgcattcacc tgccctgtgc atgtctggac 1826

tcacggagct cacccatgtg cacaagtgtg cacagtaaac gtgtttgtgg tcaacaga 1884

5 <210> 3

<211> 144

<212> PRT

<213> Homo sapiens

10 <400> 3

Met Pro Arg Gly Trp Ala Ala Pro Leu Leu Leu
1 5 10

Leu Leu Leu Gln Gly Gly Trp Gly Cys Pro Asp Leu Val Cys Tyr Thr
15 15 20 25

Asp Tyr Leu Gln Thr Val Ile Cys Ile Leu Glu Met Trp Asn Leu His
30 35 40

Pro Ser Thr Leu Thr Leu Thr Trp Gln Asp Gln Tyr Glu Glu Leu Lys
45 50 55

Asp Glu Ala Thr Ser Cys Ser Leu His Arg Ser Ala His Asn Ala Thr
60 65 70 75

His Ala Thr Tyr Thr Cys His Met Asp Val Phe His Phe Met Ala Asp
80 85 90

Asp Ile Phe Ser Val Asn Ile Thr Asp Gln Ser Gly Asn Tyr Ser Gln
95 100 105

Glu Cys Gly Ser Phe Leu Leu Ala Glu Ser Lys Ser Glu Glu Lys Ala
110 115 120

Asp Leu Ser Gly Leu Lys Lys Cys Leu Pro Pro Pro Pro Gly Val Pro
125 130 135

Gln Arg Leu Glu Leu

140

5 <210> 4

<211> 1729

<212> DNA

<213> Homo sapiens

10 <220>

<221> CDS

<222> (441).. (872)

<400> 4

15 ggcagccagc ggcctcagac agaccactg gcgtctctct gctgagtgac cgtaagctcg 60

gcgtctggcc ctctgcctgc ctctccctga gtgtggctga cagccacgca gctgtgtctg 120

tctgtctgcg gcccgatgc cctgctgcg gccgcctggg accttccttg ccgtctcttt 180

20 cctctgtctg ctgctctgtg ggacacctgc ctggaggccc agctgcccgt catcagagtg 240

acaggtctta tgacagcctg attgggtgact cgggctgggt gtggattctc accccaggcc 300

25 tctgcctgct ttctcagacc ctcatctgtc acccccacgc tgaaccacgc tgccaccccc 360

agaagcccat cagactgccc ccagcacacg gaatggattt ctgagaaaga agccgaaaca 420

gaaggcccgt gggagtcagc atg ccg cgt ggc tgg gcc gcc ccc ttg ctc ctg 473

30 Met Pro Arg Gly Trp Ala Ala Pro Leu Leu Leu

1

5

10

ctg ctg ctc cag gga ggc tgg ggc tgc ccc gac ctc gtc tgc tac acc 521

Leu Leu Leu Gln Gly Gly Trp Gly Cys Pro Asp Leu Val Cys Tyr Thr

35

15

20

25

	gat tac ctc cag acg gtc atc tgc atc ctg gaa atg tgg aac ctc cac	569
	Asp Tyr Leu Gln Thr Val Ile Cys Ile Leu Glu Met Trp Asn Leu His	
	30 35 40	
5	ccc agc acg ctc acc ctt acc tgg caa gac cag tat gaa gag ctg aag	617
	Pro Ser Thr Leu Thr Leu Thr Trp Gln Asp Gln Tyr Glu Glu Leu Lys	
	45 50 55	
10	GAC GAG GCC ACC TCC TGC AGC CTC CAC AGG TCG GCC CAC AAT GCC ACG	665
	Asp Glu Ala Thr Ser Cys Ser Leu His Arg Ser Ala His Asn Ala Thr	
	60 65 70 75	
15	CAT GCC ACC TAC ACC TGC CAC ATG GAT GTA TTC CAC TTC ATG GCC GAC	713
	His Ala Thr Tyr Thr Cys His Met Asp Val Phe His Phe Met Ala Asp	
	80 85 90	
20	GAC ATT TTC AGT GTC AAC ATC ACA GAC CAG TCT GGC AAC TAC TCC CAG	761
	Asp Ile Phe Ser Val Asn Ile Thr Asp Gln Ser Gly Asn Tyr Ser Gln	
	95 100 105	
25	GAG TGT GGC AGC TTT CTC CTG GCT GAG AGC AAG TCC GAG GAG AAA GCT	809
	Glu Cys Gly Ser Phe Leu Leu Ala Glu Ser Lys Ser Glu Glu Lys Ala	
	110 115 120	
30	gat ctc agt gga ctc aag aag tgt ctc cct cct ccc cct gga gtt ccg	857
	Asp Leu Ser Gly Leu Lys Lys Cys Leu Pro Pro Pro Pro Gly Val Pro	
	125 130 135	
35	caa aga ctc gag cta tgagctgcag gtgcgggcag ggcccatgcc tggctcctcc	912
	Gln Arg Leu Glu Leu	
	140	
	taccagggga cctggagtga atggagtgc cgggtcatct ttcagacca gtcagaggag	972
35	ttaaaggaag gctggaaccc tcacctgctg cttctcctcc tgcttgcatt agtcattcatt	1032

cctgccttct ggagcctgaa gaccatcca ttgtggaggc tatggaagaa gatatgggcc 1092
 gtccccagcc ctgagcggtt cttcatgccc ctgtacaagg gctgcagcgg agacttcaag 1152
 5 aaatgggtgg gtgcaccctt cactggctcc agcctggagc tgggaccctg gagcccagag 1212
 gtgcctcca ccctggaggt gtacagctgc caccaccca gcagccctgt ggagtgtgac 1272
 ttcaccagcc ccggggacga aggaccccc cggagctacc tccgccagtg ggtggtcatt 1332
 10 cctccgccac ttctgagccc tggaccccag gccagctaata gaggtgact ggatgtccag 1392
 agctggccag gccactgggc cctgagccag agacaaggta acctgggctg tgatgtgaag 1452
 15 acacctgcag cctttggtct cctggatggg cctttgagcc tgatgtttac agtgtctgtg 1512
 tgtgtgtgca tatgtgtgtg tgtgcatatg catgtgtgtg tgtgtgtgtg tcttaggtgc 1572
 gcagtggcat gtccacgtgt gtgtgattgc acgtgcctgt gggcctggga taatgcccata 1632
 20 ggtactccat gcattcacct gccctgtgca tgtctggact cacggagctc acccatgtgc 1692
 acaagtgtgc acagtaaacg tgtttgtggt caacaga 1729
 25 <210> 5
 <211> 237
 <212> PRT
 <213> Homo sapiens
 30 <400> 5
 Met Pro Arg Met Pro Pro Thr Pro Ala Thr Trp Met Tyr Ser Thr Ser
 1 5 10 15
 Trp Pro Thr Thr Phe Ser Val Ser Thr Ser Gln Thr Ser Leu Ala Thr
 35 20 25 30

	Thr	Pro	Arg	Ser	Val	Ala	Ala	Phe	Ser	Trp	Leu	Arg	Ala	Ser	Pro	Arg	
	35					40					45						
5	Arg	Lys	Leu	Ile	Ser	Val	Asp	Ser	Arg	Ser	Val	Ser	Leu	Leu	Pro	Leu	
	50					55					60						
	Glu	Phe	Arg	Lys	Asp	Ser	Ser	Tyr	Glu	Leu	Gln	Val	Arg	Ala	Gly	Pro	
	65					70					75					80	
10	Met	Pro	Gly	Ser	Ser	Tyr	Gln	Gly	Thr	Trp	Ser	Glu	Trp	Ser	Asp	Pro	
	85					90					95						
	Val	Ile	Phe	Gln	Thr	Gln	Ser	Glu	Glu	Leu	Lys	Glu	Gly	Trp	Asn	Pro	
15	100					105					110						
	His	Leu	Leu	Leu	Leu	Leu	Leu	Leu	Val	Ile	Val	Phe	Ile	Pro	Ala	Phe	
	115					120					125						
	Trp	Ser	Leu	Lys	Thr	His	Pro	Leu	Trp	Arg	Leu	Trp	Lys	Lys	Ile	Trp	
20	130					135					140						
	Ala	Val	Pro	Ser	Pro	Glu	Arg	Phe	Phe	Met	Pro	Leu	Tyr	Lys	Gly	Cys	
	145					150					155					160	
25	Ser	Gly	Asp	Phe	Lys	Lys	Trp	Val	Gly	Ala	Pro	Phe	Thr	Gly	Ser	Ser	
	165					170					175						
	Leu	Glu	Leu	Gly	Pro	Trp	Ser	Pro	Glu	Val	Pro	Ser	Thr	Leu	Glu	Val	
30	180					185					190						
	Tyr	Ser	Cys	His	Pro	Pro	Ser	Ser	Pro	Val	Glu	Cys	Asp	Phe	Thr	Ser	
	195					200					205						
	Pro	Gly	Asp	Glu	Gly	Pro	Pro	Arg	Ser	Tyr	Leu	Arg	Gln	Trp	Val	Val	
35	210					215					220						

Ile Pro Pro Pro Leu Ser Ser Pro Gly Pro Gln Ala Ser
 225 230 235

<210> 6

5 <211> 1729

<212> DNA

<213> Homo sapiens

<220>

10 <221> CDS

<222> (659).. (1368)

<400> 6

15 ggcagccagc ggccacagac agaccactg gcgtctctct gctgagtgac cgtaagctcg 60
 gcgtctggcc ctctgcctgc ctctccctga gtgtggctga cagccacgca gctgtgtctg 120
 tctgtctgcg gcccgatcat cctgtctgcg gccgcctggt accttccttg ccgtctcttt 180
 20 cctctgtctg ctgctctgtg ggacacctgc ctggaggccc agctgcccgt catcagagtg 240
 acaggtctta tgacagcctg attgggtgact cgggctgggt gtggattctc accccaggcc 300
 tctgcctgct ttctcagacc ctcatctgtc acccccacgc tgaaccacgc tgccaccccc 360
 25 agaagcccat cagactgccc ccagcacacg gaatggattt ctgagaaaga agccgaaaca 420
 gaaggcccgt gggagtcagc atgccgcgtg gctgggccgc ccccttgctc ctgctgctgc 480
 30 tccagggagg ctggggctgc cccgacctcg tctgctacac cgattacctc cagacgggtca 540
 tctgcatcct ggaaatgtgg aacctccacc ccagcacgct cacccttacc tggcaagacc 600
 agtatgaaga gctgaaggac gaggccacct cctgcagcct ccacaggtcg gccacaaa 658
 35 atg cca cgc atg cca cct aca cct gcc aca tgg atg tat tcc act tca 705

	Met	Pro	Arg	Met	Pro	Pro	Thr	Pro	Ala	Thr	Trp	Met	Tyr	Ser	Thr	Ser	
	1				5					10						15	
	tgg	cgc	acg	aca	ttt	tca	gtg	tca	aca	tca	cag	acc	agt	ctg	gca	act	753
5	Trp	Pro	Thr	Thr	Phe	Ser	Val	Ser	Thr	Ser	Gln	Thr	Ser	Leu	Ala	Thr	
				20					25					30			
	act	ccc	agg	agt	gtg	gca	gct	ttc	tcc	tgg	ctg	aga	gca	agt	ccg	agg	801
	Thr	Pro	Arg	Ser	Val	Ala	Ala	Phe	Ser	Trp	Leu	Arg	Ala	Ser	Pro	Arg	
10			35					40					45				
	aga	aag	ctg	atc	tca	gtg	gac	tca	aga	agt	gtc	tcc	ctc	ctc	ccc	ctg	849
	Arg	Lys	Leu	Ile	Ser	Val	Asp	Ser	Arg	Ser	Val	Ser	Leu	Leu	Pro	Leu	
		50					55				60						
15																	
	gag	ttc	cgc	aaa	gac	tgc	agc	tat	gag	ctg	cag	gtg	cgg	gca	ggg	ccc	897
	Glu	Phe	Arg	Lys	Asp	Ser	Ser	Tyr	Glu	Leu	Gln	Val	Arg	Ala	Gly	Pro	
	65					70					75				80		
20	atg	cct	ggc	tcc	tcc	tac	cag	ggg	acc	tgg	agt	gaa	tgg	agt	gac	ccg	945
	Met	Pro	Gly	Ser	Ser	Tyr	Gln	Gly	Thr	Trp	Ser	Glu	Trp	Ser	Asp	Pro	
					85					90					95		
	gtc	atc	ttt	cag	acc	cag	tca	gag	gag	tta	aag	gaa	ggc	tgg	aac	cct	993
25	Val	Ile	Phe	Gln	Thr	Gln	Ser	Glu	Glu	Leu	Lys	Glu	Gly	Trp	Asn	Pro	
				100					105					110			
	cac	ctg	ctg	ctt	ctc	ctc	ctg	ctt	gtc	ata	gtc	ttc	att	cct	gcc	ttc	1041
	His	Leu	Leu	Leu	Leu	Leu	Leu	Leu	Val	Ile	Val	Phe	Ile	Pro	Ala	Phe	
30			115					120					125				
	tgg	agc	ctg	aag	acc	cat	cca	ttg	tgg	agg	cta	tgg	aag	aag	ata	tgg	1089
	Trp	Ser	Leu	Lys	Thr	His	Pro	Leu	Trp	Arg	Leu	Trp	Lys	Lys	Ile	Trp	
		130					135				140						
35																	
	gcc	gtc	ccc	agc	cct	gag	cgg	ttc	ttc	atg	ccc	ctg	tac	aag	ggc	tgc	1137

Ala Val Pro Ser Pro Glu Arg Phe Phe Met Pro Leu Tyr Lys Gly Cys
 145 150 155 160

agc gga gac ttc aag aaa tgg gtg ggt gca ccc ttc act ggc tcc agc 1185
 5 Ser Gly Asp Phe Lys Lys Trp Val Gly Ala Pro Phe Thr Gly Ser Ser
 165 170 175

ctg gag ctg gga ccc tgg agc cca gag gtg ccc tcc acc ctg gag gtg 1233
 10 Leu Glu Leu Gly Pro Trp Ser Pro Glu Val Pro Ser Thr Leu Glu Val
 180 185 190

tac agc tgc cac cca ccc agc agc cct gtg gag tgt gac ttc acc agc 1281
 Tyr Ser Cys His Pro Pro Ser Ser Pro Val Glu Cys Asp Phe Thr Ser
 195 200 205

15 ccc ggg gac gaa gga ccc ccc cgg agc tac ctc cgc cag tgg gtg gtc 1329
 Pro Gly Asp Glu Gly Pro Pro Arg Ser Tyr Leu Arg Gln Trp Val Val
 210 215 220

20 att cct ccg cca ctt tcg agc cct gga ccc cag gcc agc taatgaggct 1378
 Ile Pro Pro Pro Leu Ser Ser Pro Gly Pro Gln Ala Ser
 225 230 235

gactggatgt ccagagctgg ccaggccact gggccctgag ccagagacaa ggtcacctgg 1438
 25 gctgtgatgt gaagacacct gcagcctttg gtctcctgga tgggcctttg agcctgatgt 1498

ttacagtgtc tgttgtgtgtg tgcatatgtg tgttgtgtgca tatgcatgtg tgttgtgtgtg 1558

30 tgtgtcttag gtgcgcagtg gcatgtccac gtgtgtgtga ttgcacgtgc ctgtgggcct 1618

gggataatgc ccatggtact ccatgcattc acctgccctg tgcatgtctg gactcacgga 1678

gctcacccat gtgcacaagt gtgcacagta aacgtgtttg tgggtcaacaga 1729
 35

<210> 7

<211> 538

<212> PRT

<213> Homo sapiens

5 <400> 7

Met Pro Arg Gly Trp Ala Ala Pro Leu Leu Leu
 1 5 10

Leu Leu Leu Gln Gly Gly Trp Gly Cys Pro Asp Leu Val Cys Tyr Thr
 10 15 20 25

Asp Tyr Leu Gln Thr Val Ile Cys Ile Leu Glu Met Trp Asn Leu His
 30 35 40

15 Pro Ser Thr Leu Thr Leu Thr Trp Gln Asp Gln Tyr Glu Glu Leu Lys
 45 50 55

Asp Glu Ala Thr Ser Cys Ser Leu His Arg Ser Ala His Asn Ala Thr
 60 65 70 75

20 His Ala Thr Tyr Thr Cys His Met Asp Val Phe His Phe Met Ala Asp
 80 85 90

25 Asp Ile Phe Ser Val Asn Ile Thr Asp Gln Ser Gly Asn Tyr Ser Gln
 95 100 105

Glu Cys Gly Ser Phe Leu Leu Ala Glu Ser Ile Lys Pro Ala Pro Pro
 110 115 120

30 Phe Asn Val Thr Val Thr Phe Ser Gly Gln Tyr Asn Ile Ser Trp Arg
 125 130 135

Ser Asp Tyr Glu Asp Pro Ala Phe Tyr Met Leu Lys Gly Lys Leu Gln
 140 145 150 155

35 Tyr Glu Leu Gln Tyr Arg Asn Arg Gly Asp Pro Trp Ala Val Ser Pro

	160	165	170
	Arg Arg Lys Leu Ile Ser Val Asp Ser Arg Ser Val Ser Leu Leu Pro		
	175	180	185
5	Leu Glu Phe Arg Lys Asp Ser Ser Tyr Glu Leu Gln Val Arg Ala Gly		
	190	195	200
	Pro Met Pro Gly Ser Ser Tyr Gln Gly Thr Trp Ser Glu Trp Ser Asp		
10	205	210	215
	Pro Val Ile Phe Gln Thr Gln Ser Glu Glu Leu Lys Glu Gly Trp Asn		
	220	225	230
15	Pro His Leu Leu Leu Leu Leu Leu Leu Val Ile Val Phe Ile Pro Ala		
	240	245	250
	Phe Trp Ser Leu Lys Thr His Pro Leu Trp Arg Leu Trp Lys Lys Ile		
	255	260	265
20	Trp Ala Val Pro Ser Pro Glu Arg Phe Phe Met Pro Leu Tyr Lys Gly		
	270	275	280
	Cys Ser Gly Asp Phe Lys Lys Trp Val Gly Ala Pro Phe Thr Gly Ser		
25	285	290	295
	Ser Leu Glu Leu Gly Pro Trp Ser Pro Glu Val Pro Ser Thr Leu Glu		
	300	305	310
30	Val Tyr Ser Cys His Pro Pro Arg Ser Pro Ala Lys Arg Leu Gln Leu		
	320	325	330
	Thr Glu Leu Gln Glu Pro Ala Glu Leu Val Glu Ser Asp Gly Val Pro		
	335	340	345
35	Lys Pro Ser Phe Trp Pro Thr Ala Gln Asn Ser Gly Gly Ser Ala Tyr		

	350		355		360
	Ser Glu Glu Arg Asp Arg Pro Tyr Gly Leu Val Ser Ile Asp Thr Val				
	365		370		375
5	Thr Val Leu Asp Ala Glu Gly Pro Cys Thr Trp Pro Cys Ser Cys Glu				
	380		385		390 395
	Asp Asp Gly Tyr Pro Ala Leu Asp Leu Asp Ala Gly Leu Glu Pro Ser				
10		400		405	410
	Pro Gly Leu Glu Asp Pro Leu Leu Asp Ala Gly Thr Thr Val Leu Ser				
		415		420	425
15	Cys Gly Cys Val Ser Ala Gly Ser Pro Gly Leu Gly Gly Pro Leu Gly				
	430		435		440
	Ser Leu Leu Asp Arg Leu Lys Pro Pro Leu Ala Asp Gly Glu Asp Trp				
	445		450		455
20	Ala Gly Gly Leu Pro Trp Gly Gly Arg Ser Pro Gly Gly Val Ser Glu				
	460		465		470 475
	Ser Glu Ala Gly Ser Pro Leu Ala Gly Leu Asp Met Asp Thr Phe Asp				
25		480		485	490
	Ser Gly Phe Val Gly Ser Asp Cys Ser Ser Pro Val Glu Cys Asp Phe				
		495		500	505
30	Thr Ser Pro Gly Asp Glu Gly Pro Pro Arg Ser Tyr Leu Arg Gln Trp				
	510		515		520
	Val Val Ile Pro Pro Pro Leu Ser Ser Pro Gly Pro Gln Ala Ser				
	525		530		535
35					
	<210> 8				

<211> 2415

<212> DNA

<213> Homo sapiens

5 <220>

<221> CDS

<222> (441).. (2054)

<400> 8

10 ggcagccagc ggcctcagac agaccactg gcgtctctct gctgagtgac cgtaagctcg 60
 gcgtctggcc ctctgcctgc ctctccctga gtgtggctga cagccacgca gctgtgtctg 120
 tctgtctgcg gcccgatgc cctgtctgcg gccgctgggt accttcttg ccgtctcttt 180
 15 cctctgtctg ctgctctgtg ggacacctgc ctggaggccc agctgcccgt catcagagtg 240
 acaggtctta tgacagcctg attgggtgact cgggctgggt gtggattctc accccaggcc 300
 20 tctgcctgct ttctcagacc ctcatctgtc acccccacgc tgaaccacgc tgccaccccc 360
 agaagcccat cagactgccc ccagcacacg gaatggattt ctgagaaaga agccgaaaca 420
 gaaggcccgt gggagtcagc atg ccg cgt ggc tgg gcc gcc ccc ttg ctc ctg 473
 25 Met Pro Arg Gly Trp Ala Ala Pro Leu Leu Leu
 1 5 10
 ctg ctg ctc cag gga ggc tgg ggc tgc ccc gac ctc gtc tgc tac acc 521
 Leu Leu Leu Gln Gly Gly Trp Gly Cys Pro Asp Leu Val Cys Tyr Thr
 30 15 20 25
 gat tac ctc cag acg gtc atc tgc atc ctg gaa atg tgg aac ctc cac 569
 Asp Tyr Leu Gln Thr Val Ile Cys Ile Leu Glu Met Trp Asn Leu His
 30 35 40
 35 ccc agc acg ctc acc ctt acc tgg caa gac cag tat gaa gag ctg aag 617

	Pro	Ser	Thr	Leu	Thr	Leu	Thr	Trp	Gln	Asp	Gln	Tyr	Glu	Glu	Leu	Lys	
	45						50					55					
	gac	gag	gcc	acc	tcc	tgc	agc	ctc	cac	agg	tcg	gcc	cac	aat	gcc	acg	665
5	Asp	Glu	Ala	Thr	Ser	Cys	Ser	Leu	His	Arg	Ser	Ala	His	Asn	Ala	Thr	
	60					65				70					75		
	cat	gcc	acc	tac	acc	tgc	cac	atg	gat	gta	ttc	cac	ttc	atg	gcc	gac	713
10	His	Ala	Thr	Tyr	Thr	Cys	His	Met	Asp	Val	Phe	His	Phe	Met	Ala	Asp	
					80				85						90		
	gac	att	ttc	agt	gtc	aac	atc	aca	gac	cag	tct	ggc	aac	tac	tcc	cag	761
	Asp	Ile	Phe	Ser	Val	Asn	Ile	Thr	Asp	Gln	Ser	Gly	Asn	Tyr	Ser	Gln	
				95				100				105					
15																	
	gag	tgt	ggc	agc	ttt	ctc	ctg	gct	gag	agc	atc	aag	ccg	gct	ccc	cct	809
	Glu	Cys	Gly	Ser	Phe	Leu	Leu	Ala	Glu	Ser	Ile	Lys	Pro	Ala	Pro	Pro	
		110					115				120						
20	ttc	aac	gtg	act	gtg	acc	ttc	tca	gga	cag	tat	aat	atc	tcc	tgg	cgc	857
	Phe	Asn	Val	Thr	Val	Thr	Phe	Ser	Gly	Gln	Tyr	Asn	Ile	Ser	Trp	Arg	
		125					130				135						
	tca	gat	tac	gaa	gac	cct	gcc	ttc	tac	atg	ctg	aag	ggc	aag	ctt	cag	905
25	Ser	Asp	Tyr	Glu	Asp	Pro	Ala	Phe	Tyr	Met	Leu	Lys	Gly	Lys	Leu	Gln	
	140				145					150					155		
	tat	gag	ctg	cag	tac	agg	aac	cgg	gga	gac	ccc	tgg	gct	gtg	agt	ccg	953
	Tyr	Glu	Leu	Gln	Tyr	Arg	Asn	Arg	Gly	Asp	Pro	Trp	Ala	Val	Ser	Pro	
30					160				165					170			
	agg	aga	aag	ctg	atc	tca	gtg	gac	tca	aga	agt	gtc	tcc	ctc	ctc	ccc	1001
	Arg	Arg	Lys	Leu	Ile	Ser	Val	Asp	Ser	Arg	Ser	Val	Ser	Leu	Leu	Pro	
				175				180				185					
35																	
	ctg	gag	ttc	cgc	aaa	gac	tcg	agc	tat	gag	ctg	cag	gtg	cgg	gca	ggg	1049

	Leu	Glu	Phe	Arg	Lys	Asp	Ser	Ser	Tyr	Glu	Leu	Gln	Val	Arg	Ala	Gly	
	190				195				200								
5	ccc	atg	cct	ggc	tcc	tcc	tac	cag	ggg	acc	tgg	agt	gaa	tgg	agt	gac	1097
	Pro	Met	Pro	Gly	Ser	Ser	Tyr	Gln	Gly	Thr	Trp	Ser	Glu	Trp	Ser	Asp	
	205				210				215								
10	ccg	gtc	atc	ttt	cag	acc	cag	tca	gag	gag	tta	aag	gaa	ggc	tgg	aac	1145
	Pro	Val	Ile	Phe	Gln	Thr	Gln	Ser	Glu	Glu	Leu	Lys	Glu	Gly	Trp	Asn	
	220				225				230				235				
15	cct	cac	ctg	ctg	ctt	ctc	ctc	ctg	ctt	gtc	ata	gtc	ttc	att	cct	gcc	1193
	Pro	His	Leu	Leu	Leu	Leu	Leu	Leu	Leu	Val	Ile	Val	Phe	Ile	Pro	Ala	
	240				245				250								
20	ttc	tgg	agc	ctg	aag	acc	cat	cca	ttg	tgg	agg	cta	tgg	aag	aag	ata	1241
	Phe	Trp	Ser	Leu	Lys	Thr	His	Pro	Leu	Trp	Arg	Leu	Trp	Lys	Lys	Ile	
	255				260				265								
25	tgg	gcc	gtc	ccc	agc	cct	gag	cgg	ttc	ttc	atg	ccc	ctg	tac	aag	ggc	1289
	Trp	Ala	Val	Pro	Ser	Pro	Glu	Arg	Phe	Phe	Met	Pro	Leu	Tyr	Lys	Gly	
	270				275				280								
30	tgc	agc	gga	gac	ttc	aag	aaa	tgg	gtg	ggt	gca	ccc	ttc	act	ggc	tcc	1337
	Cys	Ser	Gly	Asp	Phe	Lys	Lys	Trp	Val	Gly	Ala	Pro	Phe	Thr	Gly	Ser	
	285				290				295								
35	agc	ctg	gag	ctg	gga	ccc	tgg	agc	cca	gag	gtg	ccc	tcc	acc	ctg	gag	1385
	Ser	Leu	Glu	Leu	Gly	Pro	Trp	Ser	Pro	Glu	Val	Pro	Ser	Thr	Leu	Glu	
	300				305				310				315				
40	gtg	tac	agc	tgc	cac	cca	cca	cgg	agc	ccg	gcc	aag	agg	ctg	cag	ctc	1433
	Val	Tyr	Ser	Cys	His	Pro	Pro	Arg	Ser	Pro	Ala	Lys	Arg	Leu	Gln	Leu	
	320				325				330								
45	acg	gag	cta	caa	gaa	cca	gca	gag	ctg	gtg	gag	tct	gac	ggt	gtg	ccc	1481

	Thr Glu Leu Gln Glu Pro Ala Glu Leu Val Glu Ser Asp Gly Val Pro	
	335 340 345	
	aag ccc agc ttc tgg ccg aca gcc cag aac tcg ggg ggc tca gct tac	1529
5	Lys Pro Ser Phe Trp Pro Thr Ala Gln Asn Ser Gly Gly Ser Ala Tyr	
	350 355 360	
	agt gag gag agg gat cgg cca tac ggc ctg gtg tcc att gac aca gtg	1577
	Ser Glu Glu Arg Asp Arg Pro Tyr Gly Leu Val Ser Ile Asp Thr Val	
10	365 370 375	
	act gtg cta gat gca gag ggg cca tgc acc tgg ccc tgc agc tgt gag	1625
	Thr Val Leu Asp Ala Glu Gly Pro Cys Thr Trp Pro Cys Ser Cys Glu	
	380 385 390 395	
15		
	gat gac ggc tac cca gcc ctg gac ctg gat gct ggc ctg gag ccc agc	1673
	Asp Asp Gly Tyr Pro Ala Leu Asp Leu Asp Ala Gly Leu Glu Pro Ser	
	400 405 410	
	cca ggc cta gag gac cca ctc ttg gat gca ggg acc aca gtc ctg tcc	1721
20	Pro Gly Leu Glu Asp Pro Leu Leu Asp Ala Gly Thr Thr Val Leu Ser	
	415 420 425	
	tgt ggc tgt gtc tca gct ggc agc cct ggg cta gga ggg ccc ctg gga	1769
25	Cys Gly Cys Val Ser Ala Gly Ser Pro Gly Leu Gly Gly Pro Leu Gly	
	430 435 440	
	agc ctc ctg gac aga cta aag cca ccc ctt gca gat ggg gag gac tgg	1817
	Ser Leu Leu Asp Arg Leu Lys Pro Pro Leu Ala Asp Gly Glu Asp Trp	
30	445 450 455	
	gct ggg gga ctg ccc tgg ggt ggc cgg tca cct gga ggg gtc tca gag	1865
	Ala Gly Gly Leu Pro Trp Gly Gly Arg Ser Pro Gly Gly Val Ser Glu	
	460 465 470 475	
35		
	agt gag gcg ggc tca ccc ctg gcc ggc ctg gat atg gac acg ttt gac	1913

Ser Glu Ala Gly Ser Pro Leu Ala Gly Leu Asp Met Asp Thr Phe Asp
 480 485 490

agt ggc ttt gtg ggc tct gac tgc agc agc cct gtg gag tgt gac ttc 1961
 5 Ser Gly Phe Val Gly Ser Asp Cys Ser Ser Pro Val Glu Cys Asp Phe
 495 500 505

acc agc ccc ggg gac gaa gga ccc ccc cgg agc tac ctc cgc cag tgg 2009
 Thr Ser Pro Gly Asp Glu Gly Pro Pro Arg Ser Tyr Leu Arg Gln Trp
 10 510 515 520

gtg gtc att cct ccg cca ctt tgc agc cct gga ccc cag gcc agc taa 2057
 Val Val Ile Pro Pro Pro Leu Ser Ser Pro Gly Pro Gln Ala Ser
 525 530 535

15 tgaggctgac tggatgtcca gagctggcca ggccactggg ccctgagcca gagacaaggt 2117

cacctgggct gtgatgtgaa gacacctgca gcctttggtc tcctggatgg gcctttgagc 2177

20 ctgatgttta cagtgtctgt gtgtgtgtgc atatgtgtgt gtgtgcatat gcatgtgtgt 2237

gtgtgtgtgt gtcttaggtg cgcagtgga tgtccacgtg tgtgtgattg cacgtgcctg 2297

tgggcctggg ataatgcca tggtaactcca tgcattcacc tgccctgtgc atgtctggac 2357

25 tcacggagct cacccatgtg cacaagtgtg cacagtaaac gtgtttgtgg tcaacaga 2415

<210> 9
 <211> 30
 30 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Artificially Synthesized Primer Sequence

35 <400> 9

ccggctcccc ctttcaacgt gactgtgacc

30

<210> 10

<211> 30

5 <212> DNA

<213> Artificial Sequence

<220>

<223> Artificially Synthesized Primer Sequence

10

<400> 10

ggcaagcttc agtatgagct gcagtacagg

30

<210> 11

15 <211> 30

<212> DNA

<213> Artificial Sequence

<220>

20 <223> Artificially Synthesized Primer Sequence

<400> 11

accctctgac tgggtctgaa agatgaccgg

30

25 <210> 12

<211> 30

<212> DNA

<213> Artificial Sequence

30 <220>

<223> Artificially Synthesized Primer Sequence

<400> 12

catgggccct gcccgcacct gcagctcata

30

35

<210> 13

<211> 1128

<212> DNA

<213> Homo sapiens

5 <220>

<221> CDS

<222> (1)..(1125)

<400> 13

```

10  atg ccg cgt ggc tgg gcc gcc ccc ttg ctc ctg ctg ctg ctc cag gga 48
    Met Pro Arg Gly Trp Ala Ala Pro Leu Leu Leu Leu Leu Leu Gln Gly
      1             5             10             15

    ggc tgg ggc tgc ccc gac ctc gtc tgc tac acc gat tac ctc cag acg 96
15  Gly Trp Gly Cys Pro Asp Leu Val Cys Tyr Thr Asp Tyr Leu Gln Thr
      20             25             30

    gtc atc tgc atc ctg gaa atg tgg aac ctc cac ccc agc acg ctc acc 144
    Val Ile Cys Ile Leu Glu Met Trp Asn Leu His Pro Ser Thr Leu Thr
20             35             40             45

    ctt acc tgg caa gac cag tat gaa gag ctg aag gac gag gcc acc tcc 192
    Leu Thr Trp Gln Asp Gln Tyr Glu Glu Leu Lys Asp Glu Ala Thr Ser
      50             55             60

25  tgc agc ctc cac agg tcg gcc cac aat gcc acg cat gcc acc tac acc 240
    Cys Ser Leu His Arg Ser Ala His Asn Ala Thr His Ala Thr Tyr Thr
      65             70             75             80

30  tgc cac atg gat gta ttc cac ttc atg gcc gac gac att ttc agt gtc 288
    Cys His Met Asp Val Phe His Phe Met Ala Asp Asp Ile Phe Ser Val
      85             90             95

    aac atc aca gac cag tct ggc aac tac tcc cag gag tgt ggc agc ttt 336
35  Asn Ile Thr Asp Gln Ser Gly Asn Tyr Ser Gln Glu Cys Gly Ser Phe
      100            105            110

```

	ctc ctg gct gag agc atc aag ccg gct ccc cct ttc aac gtg act gtg	384
	Leu Leu Ala Glu Ser Ile Lys Pro Ala Pro Pro Phe Asn Val Thr Val	
	115 120 125	
5	acc ttc tca gga cag tat aat atc tcc tgg cgc tca gat tac gaa gac	432
	Thr Phe Ser Gly Gln Tyr Asn Ile Ser Trp Arg Ser Asp Tyr Glu Asp	
	130 135 140	
10	cct gcc ttc tac atg ctg aag ggc aag ctt cag tat gag ctg cag tac	480
	Pro Ala Phe Tyr Met Leu Lys Gly Lys Leu Gln Tyr Glu Leu Gln Tyr	
	145 150 155 160	
	agg aac cgg gga gac ccc tgg gct gtg agt ccg agg aga aag ctg atc	528
15	Arg Asn Arg Gly Asp Pro Trp Ala Val Ser Pro Arg Arg Lys Leu Ile	
	165 170 175	
	tca gtg gac tca aga agt gtc tcc ctc ctc ccc ctg gag ttc cgc aaa	576
	Ser Val Asp Ser Arg Ser Val Ser Leu Leu Pro Leu Glu Phe Arg Lys	
20	180 185 190	
	gac tcg agc tat gag ctg cag gtg cgg gca ggg ccc atg cct ggc tcc	624
	Asp Ser Ser Tyr Glu Leu Gln Val Arg Ala Gly Pro Met Pro Gly Ser	
	195 200 205	
25	tcc tac cag ggg acc tgg agt gaa tgg agt gac ccg gtc atc ttt cag	672
	Ser Tyr Gln Gly Thr Trp Ser Glu Trp Ser Asp Pro Val Ile Phe Gln	
	210 215 220	
30	acc cag tca gag acc gcc tgg atc tcc ttg gtg acc gct ctg cat cta	720
	Thr Gln Ser Glu Thr Ala Trp Ile Ser Leu Val Thr Ala Leu His Leu	
	225 230 235 240	
	gtg ctg ggc ctc agc gcc gtc ctg ggc ctg ctg ctg ctg agg tgg cag	768
35	Val Leu Gly Leu Ser Ala Val Leu Gly Leu Leu Leu Leu Arg Trp Gln	
	245 250 255	

	ttt cct gca cac tac agg aga ctg agg cat gcc ctg tgg ccc tca ctt	816
	Phe Pro Ala His Tyr Arg Arg Leu Arg His Ala Leu Trp Pro Ser Leu	
	260 265 270	
5	cca gac ctg cac cgg gtc cta ggc cag tac ctt agg gac act gca gcc	864
	Pro Asp Leu His Arg Val Leu Gly Gln Tyr Leu Arg Asp Thr Ala Ala	
	275 280 285	
10	ctg agc ccg ccc aag gcc aca gtc tca gat acc tgt gaa gaa gtg gaa	912
	Leu Ser Pro Pro Lys Ala Thr Val Ser Asp Thr Cys Glu Glu Val Glu	
	290 295 300	
	ccc agc ctc ctt gaa atc ctc ccc aag tcc tca gag agg act cct ttg	960
15	Pro Ser Leu Leu Glu Ile Leu Pro Lys Ser Ser Glu Arg Thr Pro Leu	
	305 310 315 320	
	ccc ctg tgt tcc tcc cag gcc cag atg gac tac cga aga ttg cag cct	1008
	Pro Leu Cys Ser Ser Gln Ala Gln Met Asp Tyr Arg Arg Leu Gln Pro	
20	325 330 335	
	tct tgc ctg ggg acc atg ccc ctg tct gtg tgc cca ccc atg gct gag	1056
	Ser Cys Leu Gly Thr Met Pro Leu Ser Val Cys Pro Pro Met Ala Glu	
	340 345 350	
25	tca ggg tcc tgc tgt acc acc cac att gcc aac cat tcc tac cta cca	1104
	Ser Gly Ser Cys Cys Thr Thr His Ile Ala Asn His Ser Tyr Leu Pro	
	355 360 365	
30	cta agc tat tgg cag cag cct tga	1128
	Leu Ser Tyr Trp Gln Gln Pro	
	370 375	
35	<210> 14	
	<211> 375	

<212> PRT

<213> Homo sapiens

<400> 14

```

5  Met Pro Arg Gly Trp Ala Ala Pro Leu Leu Leu Leu Leu Leu Gln Gly
    1             5             10             15

Gly Trp Gly Cys Pro Asp Leu Val Cys Tyr Thr Asp Tyr Leu Gln Thr
    20             25             30

10 Val Ile Cys Ile Leu Glu Met Trp Asn Leu His Pro Ser Thr Leu Thr
    35             40             45

Leu Thr Trp Gln Asp Gln Tyr Glu Glu Leu Lys Asp Glu Ala Thr Ser
15     50             55             60

Cys Ser Leu His Arg Ser Ala His Asn Ala Thr His Ala Thr Tyr Thr
    65             70             75             80

20 Cys His Met Asp Val Phe His Phe Met Ala Asp Asp Ile Phe Ser Val
    85             90             95

Asn Ile Thr Asp Gln Ser Gly Asn Tyr Ser Gln Glu Cys Gly Ser Phe
    100            105            110

25 Leu Leu Ala Glu Ser Ile Lys Pro Ala Pro Pro Phe Asn Val Thr Val
    115            120            125

Thr Phe Ser Gly Gln Tyr Asn Ile Ser Trp Arg Ser Asp Tyr Glu Asp
30     130            135            140

Pro Ala Phe Tyr Met Leu Lys Gly Lys Leu Gln Tyr Glu Leu Gln Tyr
    145            150            155            160

35 Arg Asn Arg Gly Asp Pro Trp Ala Val Ser Pro Arg Arg Lys Leu Ile
    165            170            175

```

	Ser Val Asp Ser Arg Ser Val Ser Leu Leu Pro Leu Glu Phe Arg Lys	
	180	185 190
5	Asp Ser Ser Tyr Glu Leu Gln Val Arg Ala Gly Pro Met Pro Gly Ser	
	195	200 205
	Ser Tyr Gln Gly Thr Trp Ser Glu Trp Ser Asp Pro Val Ile Phe Gln	
	210	215 220
10	Thr Gln Ser Glu Thr Ala Trp Ile Ser Leu Val Thr Ala Leu His Leu	
	225	230 235 240
	Val Leu Gly Leu Ser Ala Val Leu Gly Leu Leu Leu Leu Arg Trp Gln	
15	245	250 255
	Phe Pro Ala His Tyr Arg Arg Leu Arg His Ala Leu Trp Pro Ser Leu	
	260	265 270
20	Pro Asp Leu His Arg Val Leu Gly Gln Tyr Leu Arg Asp Thr Ala Ala	
	275	280 285
	Leu Ser Pro Pro Lys Ala Thr Val Ser Asp Thr Cys Glu Glu Val Glu	
	290	295 300
25	Pro Ser Leu Leu Glu Ile Leu Pro Lys Ser Ser Glu Arg Thr Pro Leu	
	305	310 315 320
	Pro Leu Cys Ser Ser Gln Ala Gln Met Asp Tyr Arg Arg Leu Gln Pro	
30	325	330 335
	Ser Cys Leu Gly Thr Met Pro Leu Ser Val Cys Pro Pro Met Ala Glu	
	340	345 350
35	Ser Gly Ser Cys Cys Thr Thr His Ile Ala Asn His Ser Tyr Leu Pro	
	355	360 365

Leu Ser Tyr Trp Gln Gln Pro
 370 375

5

<210> 15
 <211> 1383
 <212> DNA
 <213> Homo sapiens

10

<220>
 <221> CDS
 <222> (1).. (1380)

15 <400> 15

atg ccg cgt ggc tgg gcc gcc ccc ttg ctc ctg ctg ctg ctc cag gga 48
 Met Pro Arg Gly Trp Ala Ala Pro Leu Leu Leu Leu Leu Gln Gly
 1 5 10 15

20 ggc tgg ggc tgc ccc gac ctc gtc tgc tac acc gat tac ctc cag acg 96
 Gly Trp Gly Cys Pro Asp Leu Val Cys Tyr Thr Asp Tyr Leu Gln Thr
 20 25 30

25 gtc atc tgc atc ctg gaa atg tgg aac ctc cac ccc agc acg ctc acc 144
 Val Ile Cys Ile Leu Glu Met Trp Asn Leu His Pro Ser Thr Leu Thr
 35 40 45

30 ctt acc tgg caa gac cag tat gaa gag ctg aag gac gag gcc acc tcc 192
 Leu Thr Trp Gln Asp Gln Tyr Glu Glu Leu Lys Asp Glu Ala Thr Ser
 50 55 60

tgc agc ctc cac agg tcg gcc cac aat gcc acg cat gcc acc tac acc 240
 Cys Ser Leu His Arg Ser Ala His Asn Ala Thr His Ala Thr Tyr Thr
 65 70 75 80

35

tgc cac atg gat gta ttc cac ttc atg gcc gac gac att ttc agt gtc 288

	Cys His Met Asp Val Phe His Phe Met Ala Asp Asp Ile Phe Ser Val	
	85 90 95	
	aac atc aca gac cag tct ggc aac tac tcc cag gag tgt ggc agc ttt	336
5	Asn Ile Thr Asp Gln Ser Gly Asn Tyr Ser Gln Glu Cys Gly Ser Phe	
	100 105 110	
	ctc ctg gct gag agc atc aag ccg gct ccc cct ttc aac gtg act gtg	384
10	Leu Leu Ala Glu Ser Ile Lys Pro Ala Pro Pro Phe Asn Val Thr Val	
	115 120 125	
	acc ttc tca gga cag tat aat atc tcc tgg cgc tca gat tac gaa gac	432
15	Thr Phe Ser Gly Gln Tyr Asn Ile Ser Trp Arg Ser Asp Tyr Glu Asp	
	130 135 140	
	cct gcc ttc tac atg ctg aag ggc aag ctt cag tat gag ctg cag tac	480
	Pro Ala Phe Tyr Met Leu Lys Gly Lys Leu Gln Tyr Glu Leu Gln Tyr	
	145 150 155 160	
20	agg aac cgg gga gac ccc tgg gct gtg agt ccg agg aga aag ctg atc	528
	Arg Asn Arg Gly Asp Pro Trp Ala Val Ser Pro Arg Arg Lys Leu Ile	
	165 170 175	
	tca gtg gac tca aga agt gtc tcc ctc ctc ccc ctg gag ttc cgc aaa	576
25	Ser Val Asp Ser Arg Ser Val Ser Leu Leu Pro Leu Glu Phe Arg Lys	
	180 185 190	
	gac tcg agc tat gag ctg cag gtg cgg gca ggg ccc atg cct ggc tcc	624
30	Asp Ser Ser Tyr Glu Leu Gln Val Arg Ala Gly Pro Met Pro Gly Ser	
	195 200 205	
	tcc tac cag ggg acc tgg agt gaa tgg agt gac ccg gtc atc ttt cag	672
	Ser Tyr Gln Gly Thr Trp Ser Glu Trp Ser Asp Pro Val Ile Phe Gln	
	210 215 220	
35	acc cag tca gag gag ccc aaa tct tgt gac aaa act cac aca tgc cca	720

	Thr	Gln	Ser	Glu	Glu	Pro	Lys	Ser	Cys	Asp	Lys	Thr	His	Thr	Cys	Pro	
	225					230					235					240	
	ccg	tgc	cca	gca	cct	gaa	ctc	ctg	ggg	gga	ccg	tca	gtc	ttc	ctc	ttc	768
5	Pro	Cys	Pro	Ala	Pro	Glu	Leu	Leu	Gly	Gly	Pro	Ser	Val	Phe	Leu	Phe	
					245					250					255		
	ccc	cca	aaa	ccc	aag	gac	acc	ctc	atg	atc	tcc	cgg	acc	cct	gag	gtc	816
10	Pro	Pro	Lys	Pro	Lys	Asp	Thr	Leu	Met	Ile	Ser	Arg	Thr	Pro	Glu	Val	
				260					265					270			
	aca	tgc	gtg	gtg	gtg	gac	gtg	agc	cac	gaa	gac	cct	gag	gtc	aag	ttc	864
	Thr	Cys	Val	Val	Val	Asp	Val	Ser	His	Glu	Asp	Pro	Glu	Val	Lys	Phe	
				275				280					285				
15	aac	tgg	tac	gtg	gac	ggc	gtg	gag	gtg	cat	aat	gcc	aag	aca	aag	ccg	912
	Asn	Trp	Tyr	Val	Asp	Gly	Val	Glu	Val	His	Asn	Ala	Lys	Thr	Lys	Pro	
		290					295					300					
20	cgg	gag	gag	cag	tac	aac	agc	acg	tac	cgg	gtg	gtc	agc	gtc	ctc	acc	960
	Arg	Glu	Glu	Gln	Tyr	Asn	Ser	Thr	Tyr	Arg	Val	Val	Ser	Val	Leu	Thr	
	305					310				315					320		
	gtc	ctg	cac	cag	gac	tgg	ctg	aat	ggc	aag	gag	tac	aag	tgc	aag	gtc	1008
25	Val	Leu	His	Gln	Asp	Trp	Leu	Asn	Gly	Lys	Glu	Tyr	Lys	Cys	Lys	Val	
					325					330					335		
	tcc	aac	aaa	gcc	ctc	cca	gcc	ccc	atc	gag	aaa	acc	atc	tcc	aaa	gcc	1056
30	Ser	Asn	Lys	Ala	Leu	Pro	Ala	Pro	Ile	Glu	Lys	Thr	Ile	Ser	Lys	Ala	
				340					345					350			
	aaa	ggg	cag	ccc	cga	gaa	cca	cag	gtg	tac	acc	ctg	ccc	cca	tcc	cgg	1104
	Lys	Gly	Gln	Pro	Arg	Glu	Pro	Gln	Val	Tyr	Thr	Leu	Pro	Pro	Ser	Arg	
				355				360				365					
35	gat	gag	ctg	acc	aag	aac	cag	gtc	agc	ctg	acc	tgc	ctg	gtc	aaa	ggc	1152

Asp Glu Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly
 370 375 380

ttc tat ccc agc gac atc gcc gtg gag tgg gag agc aat ggg cag ccg 1200
 5 Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro
 385 390 395 400

gag aac aac tac aag acc acg cct ccc gtg ctg gac tcc gac ggc tcc 1248
 10 Glu Asn Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser
 405 410 415

ttc ttc ctc tac agc aag ctc acc gtg gac aag agc agg tgg cag cag 1296
 Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln
 420 425 430

15 ggg aac gtc ttc tca tgc tcc gtg atg cat gag gct ctg cac aac cac 1344
 Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His
 435 440 445

20 tac acg cag aag agc ctc tcc ctg tct ccg ggt aaa tga 1383
 Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
 450 455 460

25 <210> 16
 <211> 460
 <212> PRT
 <213> Homo sapiens

30 <400> 16
 Met Pro Arg Gly Trp Ala Ala Pro Leu Leu Leu Leu Leu Leu Gln Gly
 1 5 10 15

35 Gly Trp Gly Cys Pro Asp Leu Val Cys Tyr Thr Asp Tyr Leu Gln Thr
 20 25 30

	Val	Ile	Cys	Ile	Leu	Glu	Met	Trp	Asn	Leu	His	Pro	Ser	Thr	Leu	Thr
	35				40				45							
5	Leu	Thr	Trp	Gln	Asp	Gln	Tyr	Glu	Glu	Leu	Lys	Asp	Glu	Ala	Thr	Ser
	50				55				60							
	Cys	Ser	Leu	His	Arg	Ser	Ala	His	Asn	Ala	Thr	His	Ala	Thr	Tyr	Thr
	65				70				75				80			
10	Cys	His	Met	Asp	Val	Phe	His	Phe	Met	Ala	Asp	Asp	Ile	Phe	Ser	Val
	85				90				95							
	Asn	Ile	Thr	Asp	Gln	Ser	Gly	Asn	Tyr	Ser	Gln	Glu	Cys	Gly	Ser	Phe
15	100				105				110							
	Leu	Leu	Ala	Glu	Ser	Ile	Lys	Pro	Ala	Pro	Pro	Phe	Asn	Val	Thr	Val
	115				120				125							
	Thr	Phe	Ser	Gly	Gln	Tyr	Asn	Ile	Ser	Trp	Arg	Ser	Asp	Tyr	Glu	Asp
20	130				135				140							
	Pro	Ala	Phe	Tyr	Met	Leu	Lys	Gly	Lys	Leu	Gln	Tyr	Glu	Leu	Gln	Tyr
	145				150				155				160			
25	Arg	Asn	Arg	Gly	Asp	Pro	Trp	Ala	Val	Ser	Pro	Arg	Arg	Lys	Leu	Ile
	165				170				175							
	Ser	Val	Asp	Ser	Arg	Ser	Val	Ser	Leu	Leu	Pro	Leu	Glu	Phe	Arg	Lys
30	180				185				190							
	Asp	Ser	Ser	Tyr	Glu	Leu	Gln	Val	Arg	Ala	Gly	Pro	Met	Pro	Gly	Ser
	195				200				205							
	Ser	Tyr	Gln	Gly	Thr	Trp	Ser	Glu	Trp	Ser	Asp	Pro	Val	Ile	Phe	Gln
35	210				215				220							

	Thr	Gln	Ser	Glu	Glu	Pro	Lys	Ser	Cys	Asp	Lys	Thr	His	Thr	Cys	Pro	
	225					230					235					240	
5	Pro	Cys	Pro	Ala	Pro	Glu	Leu	Leu	Gly	Gly	Pro	Ser	Val	Phe	Leu	Phe	
						245				250					255		
	Pro	Pro	Lys	Pro	Lys	Asp	Thr	Leu	Met	Ile	Ser	Arg	Thr	Pro	Glu	Val	
						260				265					270		
10	Thr	Cys	Val	Val	Val	Asp	Val	Ser	His	Glu	Asp	Pro	Glu	Val	Lys	Phe	
						275				280					285		
	Asn	Trp	Tyr	Val	Asp	Gly	Val	Glu	Val	His	Asn	Ala	Lys	Thr	Lys	Pro	
						290				295					300		
15	Arg	Glu	Glu	Gln	Tyr	Asn	Ser	Thr	Tyr	Arg	Val	Val	Ser	Val	Leu	Thr	
						305				310					315		
	Val	Leu	His	Gln	Asp	Trp	Leu	Asn	Gly	Lys	Glu	Tyr	Lys	Cys	Lys	Val	
20						325				330					335		
	Ser	Asn	Lys	Ala	Leu	Pro	Ala	Pro	Ile	Glu	Lys	Thr	Ile	Ser	Lys	Ala	
						340				345					350		
25	Lys	Gly	Gln	Pro	Arg	Glu	Pro	Gln	Val	Tyr	Thr	Leu	Pro	Pro	Ser	Arg	
						355				360					365		
	Asp	Glu	Leu	Thr	Lys	Asn	Gln	Val	Ser	Leu	Thr	Cys	Leu	Val	Lys	Gly	
						370				375					380		
30	Phe	Tyr	Pro	Ser	Asp	Ile	Ala	Val	Glu	Trp	Glu	Ser	Asn	Gly	Gln	Pro	
						385				390					395		
	Glu	Asn	Asn	Tyr	Lys	Thr	Thr	Pro	Pro	Val	Leu	Asp	Ser	Asp	Gly	Ser	
35						405				410					415		

Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln
 420 425 430

Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His
 5 435 440 445

Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
 450 455 460

10

<210> 17
 <211> 477
 <212> DNA
 <213> Homo sapiens

15

<220>
 <221> CDS
 <222> (1).. (474)

20

<400> 17
 atg ccg cgt ggc tgg gcc gcc ccc ttg ctc ctg ctg ctg ctc cag gga 48
 Met Pro Arg Gly Trp Ala Ala Pro Leu Leu Leu Leu Leu Leu Gln Gly
 1 5 10 15

25

ggc tgg ggc tgc ccc gac ctc gtc tgc tac acc gat tac ctc cag acg 96
 Gly Trp Gly Cys Pro Asp Leu Val Cys Tyr Thr Asp Tyr Leu Gln Thr
 20 25 30

30

gtc atc tgc atc ctg gaa atg tgg aac ctc cac ccc agc acg ctc acc 144
 Val Ile Cys Ile Leu Glu Met Trp Asn Leu His Pro Ser Thr Leu Thr
 35 40 45

35

ctt acc tgg caa gac cag tat gaa gag ctg aag gac gag gcc acc tcc 192
 Leu Thr Trp Gln Asp Gln Tyr Glu Glu Leu Lys Asp Glu Ala Thr Ser
 50 55 60

tgc agc ctc cac agg tcg gcc cac aat gcc acg cat gcc acc tac acc	240
Cys Ser Leu His Arg Ser Ala His Asn Ala Thr His Ala Thr Tyr Thr	
65 70 75 80	

5 tgc cac atg gat gta ttc cac ttc atg gcc gac gac att ttc agt gtc 288
Cys His Met Asp Val Phe His Phe Met Ala Asp Asp Ile Phe Ser Val
 85 90 95

aac atc aca gac cag tct ggc aac tac tcc cag gag tgt ggc agc ttt 336
 10 Asn Ile Thr Asp Gln Ser Gly Asn Tyr Ser Gln Glu Cys Gly Ser Phe
 100 105 110

ctc ctg gct gag agc aag tcc gag gag aaa gct gat ctc agt gga ctc 384
Leu Leu Ala Glu Ser Lys Ser Glu Glu Lys Ala Asp Leu Ser Gly Leu
15 115 120 125

aag aag tgt ctc cct cct ccc cct gga gtt ccg caa aga ctc gag cta 432
Lys Lys Cys Leu Pro Pro Pro Pro Gly Val Pro Gln Arg Leu Glu Leu
130 135 140

20
agg gcg cgc cag gac tac aag gac gac gat gac aag acg cgt taa 477
Arg Ala Arg Gln Asp Tyr Lys Asp Asp Asp Asp Lys Thr Arg
145 150 155

25

<210> 18

<211> 158

<212> PRT

<213> Homo sapiens

30
 <400> 18
 Met Pro Arg Gly Trp Ala Ala Pro Leu Leu Leu Leu Leu Leu Gln Gly
 1 5 10 15

35 Gly Trp Gly Cys Pro Asp Leu Val Cys Tyr Thr Asp Tyr Leu Gln Thr
20 25 30

Val Ile Cys Ile Leu Glu Met Trp Asn Leu His Pro Ser Thr Leu Thr
 35 40 45

5 Leu Thr Trp Gln Asp Gln Tyr Glu Glu Leu Lys Asp Glu Ala Thr Ser
 50 55 60

Cys Ser Leu His Arg Ser Ala His Asn Ala Thr His Ala Thr Tyr Thr
 65 70 75 80

10 Cys His Met Asp Val Phe His Phe Met Ala Asp Asp Ile Phe Ser Val
 85 90 95

Asn Ile Thr Asp Gln Ser Gly Asn Tyr Ser Gln Glu Cys Gly Ser Phe
 15 100 105 110

Leu Leu Ala Glu Ser Lys Ser Glu Glu Lys Ala Asp Leu Ser Gly Leu
 115 120 125

20 Lys Lys Cys Leu Pro Pro Pro Pro Gly Val Pro Gln Arg Leu Glu Leu
 130 135 140

Arg Ala Arg Gln Asp Tyr Lys Asp Asp Asp Asp Lys Thr Arg
 145 150 155

25

[Brief Description of the Drawings]

[Fig. 1]

Fig. 1 is a schematic diagram showing the results of BlastX search where the query was 180 nucleotides of 40952-40966 including 40952-40966, the only probe sequence within the AC002303. "#": For only NR8 the number was indicated by the nucleotide number. The underline of the NR8 sequence shows the portion corresponding to the exon. Other underlined sequences show identical amino acids.

[Fig. 2]

35 Fig. 2 is a schematic diagram showing the results of BlastX scanning of 180 nucleotides in both the 5' and 3' directions, where the search

centered on the 180 nucleotides of 40952-40966 containing 40952-40966, the only probe sequence within the AC002303.

[Fig. 3]

Fig. 3 shows the electrophoresis results of the amplification done
5 by the RT-PCR method for the combinations of SN1/AS1, SN1/AS2, SN2/AS1, and SN2/AS2 primers using human fetal liver and skeletal muscle cDNA as templates.

[Fig. 4]

Fig. 4 shows the electrophoretic results of the 5'-RACE method
10 and 3'-RACE method using human fetal liver cDNA as the template.

[Fig. 5]

Fig. 5 shows the nucleotide sequence and the amino acid sequence
of NR8 α cDNA. The arrows show the positions of primers used for RT-PCR. They are, SN1 (798-827), SN2 (894-923), AS2 (1055-1026), and AS1
15 (1127-1098) from the 5' side, in their order. For two bases at the 5' end of AS1, AC, which is derived from the genomic sequence, was used in place of CT.

[Fig. 6]

Fig. 6 is the continuation of Fig. 5 showing the nucleotide sequence
20 and the amino acid sequence of NR8 α cDNA.

[Fig. 7]

Fig. 7 shows the nucleotide sequence and the amino acid sequence
of NR8 β cDNA. Two possible open reading frames (ORF) are shown.

[Fig. 8]

Fig. 8 is the continuation of Fig. 7 showing the nucleotide sequence
25 and the amino acid sequence of NR8 β cDNA.

[Fig. 9]

Fig. 9 shows the nucleotide sequence and the amino acid sequence
of NR8 γ cDNA. The 177 amino acids inserted by selective splicing are
30 underlined.

[Fig. 10]

Fig. 10 is the continuation of Fig. 9 showing the nucleotide sequence
and the amino acid sequence of NR8 γ cDNA. The 177 amino acids inserted
by selective splicing are underlined.

35 [Fig. 11]

Fig. 11 is the continuation of Fig. 10 showing the nucleotide

sequence and the amino acid sequence of NR8 γ cDNA.

[Fig. 12]

Fig. 12 shows the results of Northern blot analysis of NR8 expression in each organ.

5 [Fig. 13]

Fig. 13 is a schematic diagram showing the structure of the NR8 gene. Other repetitives include, (CA) n , (CAGA) n , (TGGA) n , (CATA) n , (TA) n , (GA) n , (GGAA) n , (CATG) n , (GAAA) n , MSTA, AT-rich, MLT1A1, LINE2, FLAM_C, MER63A, and MSTB.

10 [Fig. 14]

Fig. 14 is a schematic diagram showing the structure of expressible proteins constructed in the expression vector.

[Document Name] Drawings

[Fig. 1]

NR#	40862	<u>SLLPLEFRKDSSYELQVRAGMPGSSYQGTWSEWSDPVIFQIQSEGRCEAGMDTPLL</u>	41032
hTPOR	442	<u>LELRPRSRYRLQLRAR-LNGPTYQGPWSSWSDPTRVETATE</u>	481
hOBR	292	<u>SLLVDSILPGSSYEVQVRGKRLDGP---GIWSDWSTPRVFTTQ</u>	331
hIL2Rb	201	<u>DTQYEFQVRVKPLQGEFT--TWSPWSQPLAFRTK</u>	232
hIL7R	189	<u>TLLQRKLQPAAMYIEIKVRS--IPDHYFKGFWSEWSPSYFRTPEINNSSGEMDPILL</u>	243
hGM-CSFRb	196	<u>TLGPEHLMPPSSTYVARVTRLAPGSRLSGRPSKWSPEVCWDSQ</u>	238
	419	<u>TGYNGIWSEWSEARSWDIES</u>	438
mIL3Rb	200	<u>NLEPKLFLPNSIYAARVTRLASGSSLGRPSRWSPVHWSQ</u>	242
	404	<u>QLEPDTSYCARVRVKPI--SDYDGIWSEWSNEYTWIT</u>	438
hIL5Ra	302	<u>SKYDVQVRAAVSSMCREAGLWSEWSQPI</u>	329
hIL9R	241	<u>YTGQWSEWSQPVCFQ</u>	255
hEPOR	211	<u>RGRTRYTFAVRAR-MAEPSFGGFWSAWSEPVSLLTPSD</u>	247
hIL2Rr	209	<u>SLPSVDGQKRYTFRVRSRFRNPLCGSAQH--WSEWSHP</u>	244
hIL12R	197	<u>LCPLEMNVAQEFQLRRRQLGSGSS-----WSKWSSPV</u>	229
hIL12Rb	282	<u>LDLKPFTYEYEFQISSKL----HLYKGSWSDWSES LRAQTPEE</u>	319

[Fig. 2]

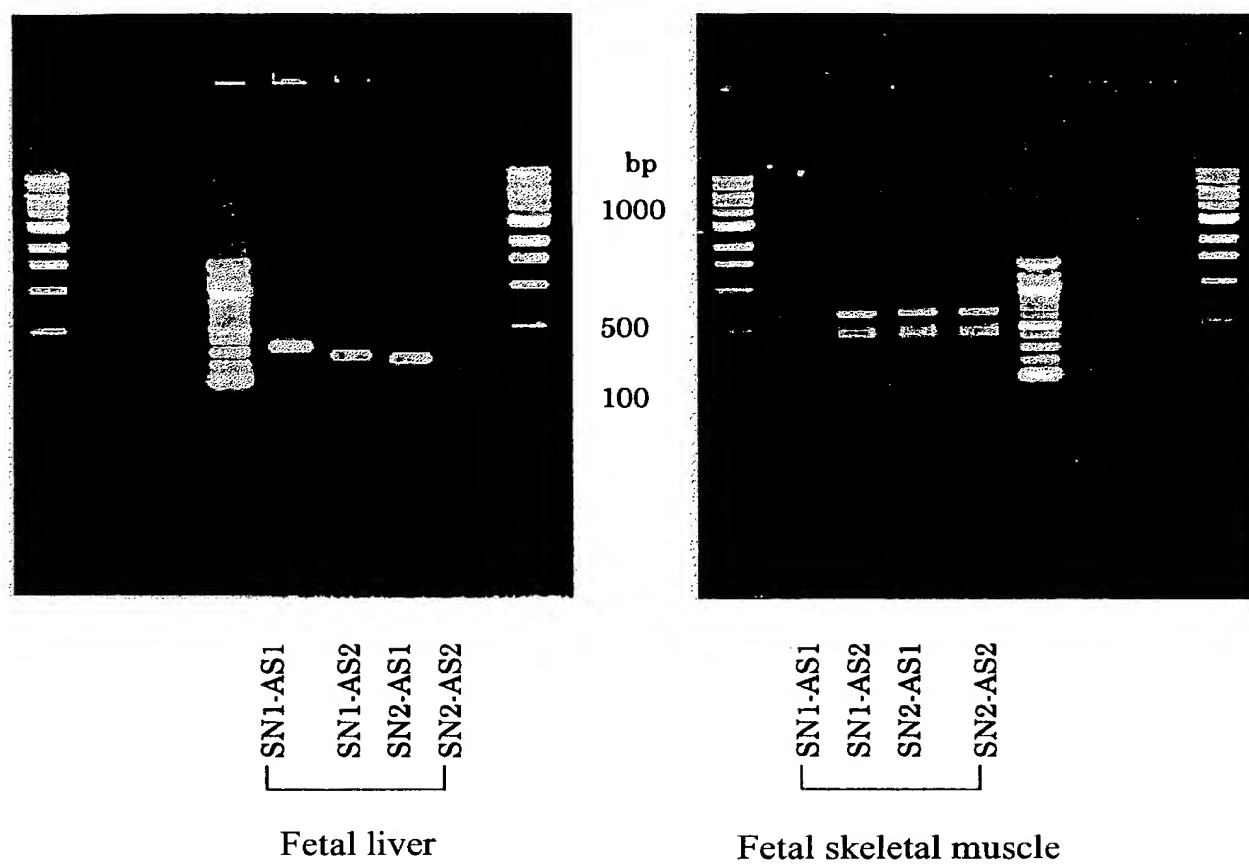
[(Query) : 39181-39360]

NR8	39233	<u>HQVKPAPPFN</u> -- <u>VTVTFSGQYNISWRS-DYEDP</u> ----- <u>AFYMLKGKLQY</u>	39355
hIL6Ra	214	<u>LQDP</u> <u>PANI</u> -- <u>TVTAVAR</u> -- <u>NPRWL</u> <u>SVTWQDPHSWNSSFYRLRFELRY</u>	257
hgp130	218	<u>YKVKPNPPHN</u> L-- <u>SVINSEELSSILKLTWT</u> - <u>NPSIKSV</u> -- <u>IILKYNIQY</u>	261
rOBRb	234	<u>VKDP</u> <u>PLGLRMEV</u> <u>TDG</u> <u>NLKISWD</u> <u>S-QTKAP</u>	263

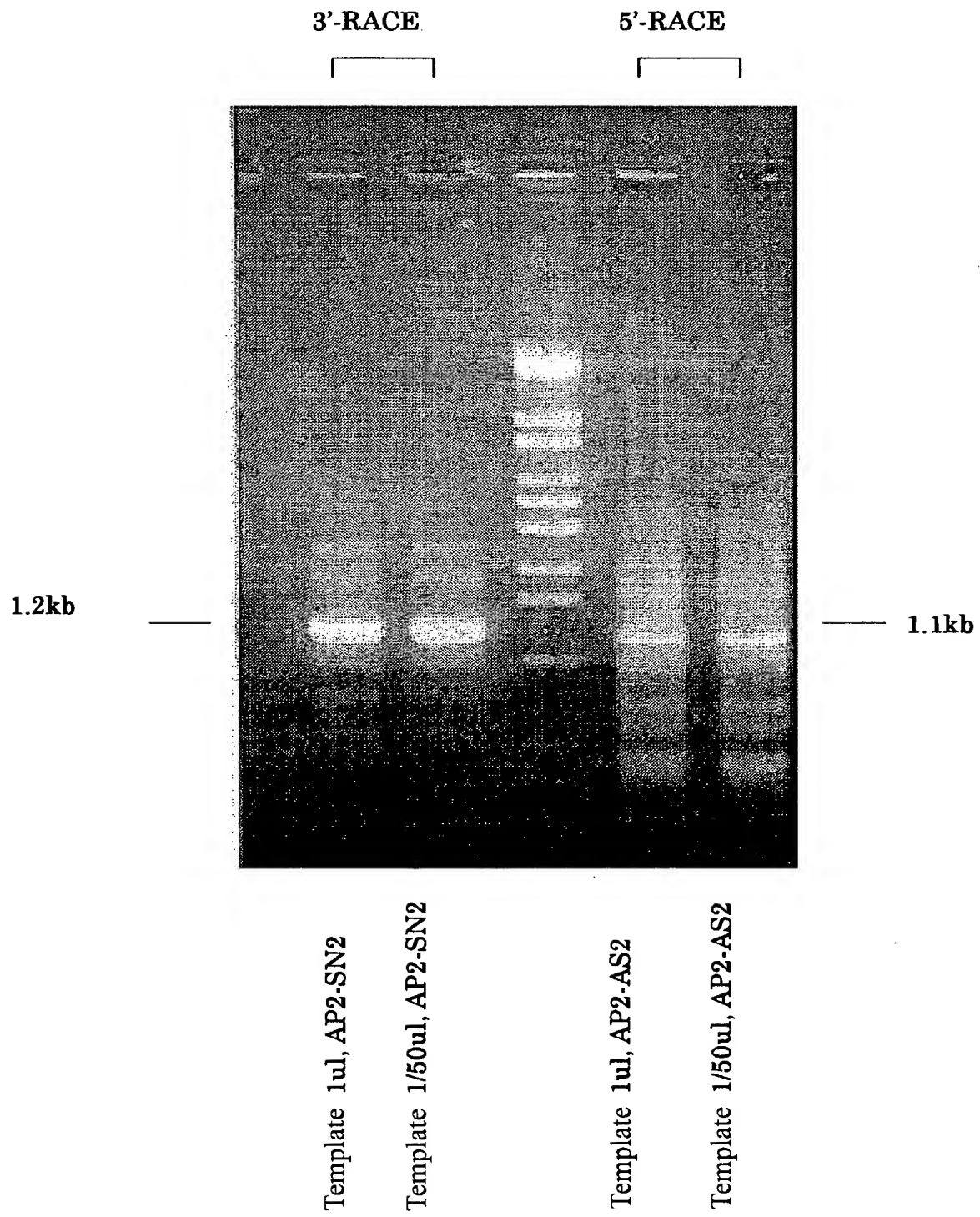
[(Query) : 42301-42480]

NR8	42307	<u>VPSPERFFMPLYKGC</u> <u>SGDFK</u>	42366
mIL9R	305	<u>IPSP</u> <u>EAF</u> <u>FHPLYSVYH</u> <u>GDFQ</u>	324
hIL9R	305	<u>VPSP</u> <u>AMFFQPLYSVH</u> <u>NGNFQ</u>	324

[Fig. 3]



[Fig. 4]



[Fig. 5]

10 20 30 40 50 60 70 80
 GGCAGCCAGCGGCCTCAGACAGACCCACTGGCGTCTCTCTGCTGAGTGACCGTAAGCTCGGCGTCTGGCCCTCTGCCTGC
 90 100 110 120 130 140 150 160
 CTCTCCCTGAGTGTGGCTGACAGCCACGCAGCTGTGTCTGTCTGTCTGCGGCCCGTGCATCCCTGCTGCGGCCGCTGGT
 170 180 190 200 210 220 230 240
 ACCTTCCTTGCCGTCTCTTTCTCTGTCTGCTGCTCTGTGGGACACCTGCCTGGAGGCCAGCTGCCCGTCATCAGAGTG
 250 260 270 280 290 300 310 320
 ACAGGTCTTATGACAGCCTGATTGGTGACTCGGGCTGGGTGTGGATTCTACCCCAGGCCTCTGCCTGCTTTCTCAGACC
 330 340 350 360 370 380 390 400
 CTCATCTGTACCCCCACGCTGAACCCAGCTGCCACCCCAGAAGCCCATCAGACTGCCCCAGCACACGGAATGGATTT
 410 420 430 440 450 460 470 480
 CTGAGAAAGAAGCCGAAACAGAAGGCCCGTGGGAGTCAGCATGCCGCGTGGCTGGGCCGCCCCCTTGCTCCTGCTGCTGC
 M P R G W A A P L L L L L L
 490 500 510 520 530 540 550 560
 TCCAGGGAGGCTGGGGCTGCCCCGACCTCGTCTGCTACACCGATTACCTCCAGACGGTCATCTGCATCCTGGAAATGTGG
 Q G G W G C P D L V C Y T D Y L Q T V I C I L E M W
 570 580 590 600 610 620 630 640
 AACCTCCACCCCAGCAGCTCACCTTACCTGGCAAGACCAGTATGAAGAGCTGAAGGACGAGGCCACCTCCTGCAGCCT
 N L H P S T L T L T W Q D Q Y E E L K D E A T S C S L
 650 660 670 680 690 700 710 720
 CCACAGGTGCGCCCAATGCCACGCATGCCACCTACACCTGCCACATGGATGTATTCCACTTCATGGCCGACGACATTT
 H R S A H N A T H A T Y T C H M D V F H F M A D D I F
 730 740 750 760 770 780 790 800
 TCAGTGTCACATCAGACAGCTCTGGCAACTACTCCCAGGAGTGTGGCAGCTTTCTCCTGGCTGAGAGCATCAAGCCG
 S V N I T D Q S G N Y S Q E C G S F L L A E S I K P
 810 820 830 840 850 860 870 880
 GCTCCCCCTTTCAACGTGACTGTGACCTTCTCAGGACAGTATAATATCTCCTGGCGCTCAGATTACGAAGACCCTGCCTT
 A P P F N V T V T F S G Q Y N I S W R S D Y E D P A F
 890 900 910 920 930 940 950 960
 CTACATGCTGAAGGGCAAGCTTCAGTATGAGCTGCAGTACAGGAACCGGGGAGACCCTGGGCTGTGAGTCCGAGGAGAA
 Y M L K G K L Q Y E L Q Y R N R G D P W A V S P R R K
 970 980 990 1000 1010 1020 1030 1040
 AGCTGATCTCAGTGGACTCAAGAAGTGTCTCCCTCCTCCCCTGGAGTTCGCAAAGACTCGAGCTATGAGCTGCAGGTG
 L I S V D S R S V S L L P L E F R K D S S Y E L Q V
 1050 1060 1070 1080 1090 1100 1110 1120
 CGGGCAGGGCCCATGCCTGGCTCCTCCTACCAGGGGACCTGGAGTGAATGGAGTGACCCGGTCATCTTTCAGACCCAGTC
 R A G P M P G S S Y Q G T W S E W S D P V I F Q T Q S

1130	1140	1150	1160	1170	1180	1190	1200																			
AGAGGAGTTAAAGGAAGGCTGGAACCCTCACCTGCTGCTTCTCCTCCTGCTTGTGCATAGTCTTCATTCCCTGCCTTCTGGA																										
E	E	L	K	E	G	W	N	P	H	L	L	L	L	L	L	V	I	V	F	I	P	A	F	W	S	
1210	1220	1230	1240	1250	1260	1270	1280																			
GCCTGAAGACCCATCCATTGTGGAGGCTATGGAAGAAGATATGGGCCGTCCCCAGCCCTGAGCGGTTCTTCATGCCCTG																										
L	K	T	H	P	L	W	R	L	W	K	K	I	W	A	V	P	S	P	E	R	F	F	M	P	L	
1290	1300	1310	1320	1330	1340	1350	1360																			
TACAAGGGCTGCAGCGGAGACTTCAAGAAATGGGTGGGTGCACCCTTCACTGGCTCCAGCCTGGAGCTGGGACCCTGGAG																										
Y	K	G	C	S	G	D	F	K	K	W	V	G	A	P	F	T	G	S	S	L	E	L	G	P	W	S
1370	1380	1390	1400	1410	1420	1430	1440																			
CCCAGAGGTGCCCTCCACCCTGGAGGTGTACAGCTGCCACCCACCCAGCAGCCCTGTGGAGTGTGACTTCACCAGCCCCG																										
P	E	V	P	S	T	L	E	V	Y	S	C	H	P	P	S	S	P	V	E	C	D	F	T	S	P	G
1450	1460	1470	1480	1490	1500	1510	1520																			
GGGACGAAGGACCCCCCGGAGCTACCTCCGCCAGTGGGTGGTCATTCTCCGCCACTTTCGAGCCCTGGACCCAGGCC																										
D	E	G	P	P	R	S	Y	L	R	Q	W	V	V	I	P	P	P	L	S	S	P	G	P	Q	A	
1530	1540	1550	1560	1570	1580	1590	1600																			
AGCTAATGAGGCTGACTGGATGTCCAGAGCTGGCCAGGCCACTGGGCCCTGAGCCAGAGACAAGGTCACCTGGGCTGTGA																										
S	*	*																								
1610	1620	1630	1640	1650	1660	1670	1680																			
TGTGAAGACACCTGCAGCCTTTGGTCTCCTGGATGGGCCTTTGAGCCTGATGTTTACAGTGTCTGTGTGTGTGTGCATAT																										
1690	1700	1710	1720	1730	1740	1750	1760																			
GTGTGTGTGTGCATATGCATGTGTGTGTGTGTGTGTGTCTTAGGTGCGCAGTGGCATGTCCACGTGTGTGTGATTGCACG																										
1770	1780	1790	1800	1810	1820	1830	1840																			
TGCCTGTGGGCCTGGGATAATGCCCATGGTACTCCATGCATTACCTGCCCTGTGCATGTCTGGACTCACGGAGCTCACC																										
1850	1860	1870	1880	1890	1900	1910	1920																			
CATGTGCACAAGTGTGCACAGTAAACGTGTTTGTGGTCAACAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA																										
1930	AAAAAAAAAAAAAA																									

[Fig. 7]

10 20 30 40 50 60 70 80
GGCAGCCAGCGGCCTCAGACAGACCCACTGGCGTCTCTCTGCTGAGTGACCGTAAGCTCGGCGTCTGGCCCTCTGCCTGC

90 100 110 120 130 140 150 160
CTCTCCCTGAGTGTGGCTGACAGCCACGCAGCTGTGTCTGTCTGTCTGCGGCCCGTGCATCCCTGCTGCGGCCGCCTGGT

170 180 190 200 210 220 230 240
ACCTTCCTTGCCGTCTCTTTCTCTGTCTGCTGCTCTGTGGGACACCTGCCTGGAGGCCAGCTGCCCGTCATCAGAGTG

250 260 270 280 290 300 310 320
ACAGGTCTTATGACAGCCTGATTGGTGACTCGGGCTGGGTGTGGATTCTACCCCAGGCCTCTGCCTGCTTTCTCAGACC

330 340 350 360 370 380 390 400
CTCATCTGTACCCCCACGCTGAACCCAGCTGCCACCCCCAGAAGCCCATCAGACTGCCCCAGCACACGGAATGGATT

410 420 430 440 450 460 470 480
CTGAGAAAGAAGCCGAAACAGAAGGCCGTGGGAGTCAGCATGCCCGTGGCTGGGCCGCCCTTGCTCTGCTGCTGC
M P R G W A A P L L L L L L L

490 500 510 520 530 540 550 560
TCCAGGGAGGCTGGGGCTGCCCCGACCTCGTCTGCTACACCGATTACCTCCAGACGGTCATCTGCATCCTGAAATGTGG
Q G G W G C P D L V C Y T D Y L Q T V I C I L E M W

570 580 590 600 610 620 630 640
AACCTCCACCCAGCACGCTCACCTTACCTGGCAAGACCAGTATGAAGAGCTGAAGGACGAGGCCACCTCCTGCAGCCT
N L H P S T L T L T W Q D Q Y E E L K D E A T S C S L

650 660 670 680 690 700 710 720
CCACAGGTGCGCCACAATGCCACGCATGCCACCTACACCTGCCACATGGATGTATTCCACTTCATGGCCGACGACATTT
H R S A H N A T H A T Y T C H M D V F H F M A D D I F
M P R M P P T P A T W M Y S T S W P T T F

730 740 750 760 770 780 790 800
TCAGTGTCAACATCACAGACCAGTCTGGCAACTACTCCAGGAGTGTGGCAGCTTTCTCCTGGCTGAGAGCAAGTCCGAG
S V N I T D Q S G N Y S Q E C G S F L L A E S K S E
S V S T S Q T S L A T T P R S V A A F S W L R A S P R

810 820 830 840 850 860 870 880
GAGAAAGCTGATCTCAGTGGACTCAAGAAGTGTCTCCCTCCTCCCCCTGGAGTTCGCAAAGACTCGAGCTATGAGCTGC
E K A D L S G L K K C L P P P P G V P Q R L E L *

890 900 910 920 930 940 950 960
R K L I S V D S R S V S L L P L E F R K D S S Y E L Q
AGGTGCGGGCAGGGCCCATGCCTGGCTCCTCCTACCAGGGGACCTGGAGTGAATGGAGTGACCCGGTCATCTTTCAGACC

V R A G P M P G S S Y Q G T W S E W S D P V I F Q T

[Fig. 8]

970 980 990 1000 1010 1020 1030 1040
CAGTCAGAGGAGTTAAAGGAAGGCTGGAACCTCACCTGCTGCTTCTCCTCCTGCTTGTGCATAGTCTTCATTCTGCCTT

Q S E E L K E G W N P H L L L L L L L V I V F I P A F
1050 1060 1070 1080 1090 1100 1110 1120
CTGGAGCCTGAAGACCCATCCATTGTGGAGGCTATGGAAGAAGATATGGGCCGTCCCAGCCCTGAGCGGTTCTTCATGC

W S L K T H P L W R L W K K I W A V P S P E R F F M P
1130 1140 1150 1160 1170 1180 1190 1200
CCCTGTACAAGGGCTGCAGCGGAGACTTCAAGAAATGGGTGGGTGCACCCTTCACTGGCTCCAGCCTGGAGCTGGGACCC

L Y K G C S G D F K K W V G A P F T G S S L E L G P
1210 1220 1230 1240 1250 1260 1270 1280
TGAGGCCAGAGGTGCCCTCCACCCTGGAGGTGTACAGCTGCCACCCACCCAGCAGCCCTGTGGAGTGTGACTTCACCAG

W S P E V P S T L E V Y S C H P P S S P V E C D F T S
1290 1300 1310 1320 1330 1340 1350 1360
CCCCGGGGACGAAGGACCCCCCGGAGCTACCTCCGCCAGTGGGTGGTCATTCTCCGCCACTTTCGAGCCCTGGACCCC

P G D E G P P R S Y L R Q W V V I P P P L S S P G P Q

1370 1380 1390 1400 1410 1420 1430 1440
AGGCCAGCTAATGAGGCTGACTGGATGTCCAGAGCTGGCCAGGCCACTGGGCCCTGAGCCAGAGACAAGGTCACCTGGGC

A S * *

1450 1460 1470 1480 1490 1500 1510 1520
TGTGATGTGAAGACACCTGCAGCCTTTGGTCTCCTGGATGGGCCTTTGAGCCTGATGTTACAGTGTCTGTGTGTGTGTG

1530 1540 1550 1560 1570 1580 1590 1600
CATATGTGTGTGTGTGCATATGCATGTGTGTGTGTGTGTGTCTTAGGTGCGCAGTGGCATGTCCACGTGTGTGTGATT

1610 1620 1630 1640 1650 1660 1670 1680
GCACGTGCCTGTGGGCCTGGGATAATGCCCATGGTACTCCATGCATTACCTGCCCTGTGCATGTCTGGACTCACGGAGC

1690 1700 1710 1720 1730 1740 1750 1760
TCACCCATGTGCACAAGTGTGCACAGTAAACGTGTTTGTGGTCAACAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

1770 1780
AAAAAAAAAAAAAAAAAAAAA

[Fig. 9]

10 20 30 40 50 60 70 80
GGCAGCCAGCGGCTCAGACAGACCCACTGGCGTCTCTGCTGAGTGACCGTAAGCTCGGCGTCTGGCCCTCTGCCTGC

90 100 110 120 130 140 150 160
CTCTCCCTGAGTGTGGCTGACAGCCACGCAGCTGTGTCTGTCTGTCTGCGGCCCGTGCATCCCTGCTGCGGCCCGCTGGT

170 180 190 200 210 220 230 240
ACCTTCCTTGCCGTCTCTTTCCTCTGTCTGCTGCTCTGTGGGACACCTGCCTGGAGGCCAGCTGCCCGTCATCAGAGTG

250 260 270 280 290 300 310 320
ACAGGTCTTATGACAGCCTGATTGGTGACTCGGGCTGGGTGTGGATTCTCACCCAGGCCTCTGCCTGCTTTCTCAGACC

330 340 350 360 370 380 390 400
CTCATCTGTACCCCCACGCTGAACCCAGCTGCCACCCCCAGAAGCCCATCAGACTGCCCCAGCACACGGAATGGATT

410 420 430 440 450 460 470 480
CTGAGAAAGAAGCCGAAACAGAAGGCCCGTGGGAGTCAGCATGCCGCGTGGCTGGGCCGCCCTTGCTCCTGCTGCTGC

490 500 510 520 530 540 550 560
M P R G W A A P L L L L L L
TCCAGGGAGGCTGGGGCTGCCCCGACCTCGTCTGCTACACCGATTACCTCCAGACGGTCATCTGCATCCTGAAATGTGG

570 580 590 600 610 620 630 640
Q G G W G C P D L V C Y T D Y L Q T V I C I L E M W
AACCTCCACCCAGCAGCTCACCTTACCTGGCAAGACCAGTATGAAGAGCTGAAGGACGAGGCCACCTCCTGCAGCCT

650 660 670 680 690 700 710 720
N L H P S T L T L T W Q D Q Y E E L K D E A T S C S L
CCACAGGTGCGCCACAATGCCACGCATGCCACCTACACCTGCCACATGGATGTATTCCACTTCATGGCCGACGACATTT

730 740 750 760 770 780 790 800
H R S A H N A T H A T Y T C H M D V F H F M A D D I F
TCAGTGTCAACATCAGACCAGTCTGGCAACTACTCCAGGAGTGTGGCAGCTTTCTCCTGGCTGAGAGCATCAAGCCG

810 820 830 840 850 860 870 880
S V N I T D Q S G N Y S Q E C G S F L L A E S I K P
GCTCCCCCTTTCAACGTGACTGTGACCTTCTCAGGACAGTATAATATCTCCTGGCGCTCAGATTACGAAGACCCTGCCTT

890 900 910 920 930 940 950 960
A P P F N V T V T F S G Q Y N I S W R S D Y E D P A F
CTACATGCTGAAGGGCAAGCTTCAGTATGAGCTGCAGTACAGGAACCGGGAGACCCCTGGGCTGTGAGTCCGAGGAGAA

970 980 990 1000 1010 1020 1030 1040
Y M L K G K L Q Y E L Q Y R N R G D P W A V S P R R K
AGCTGATCTCAGTGGACTCAAGAAGTGTCTCCCTCCTCCCCCTGGAGTTCGCAAAGACTCGAGCTATGAGCTGCAGGTG

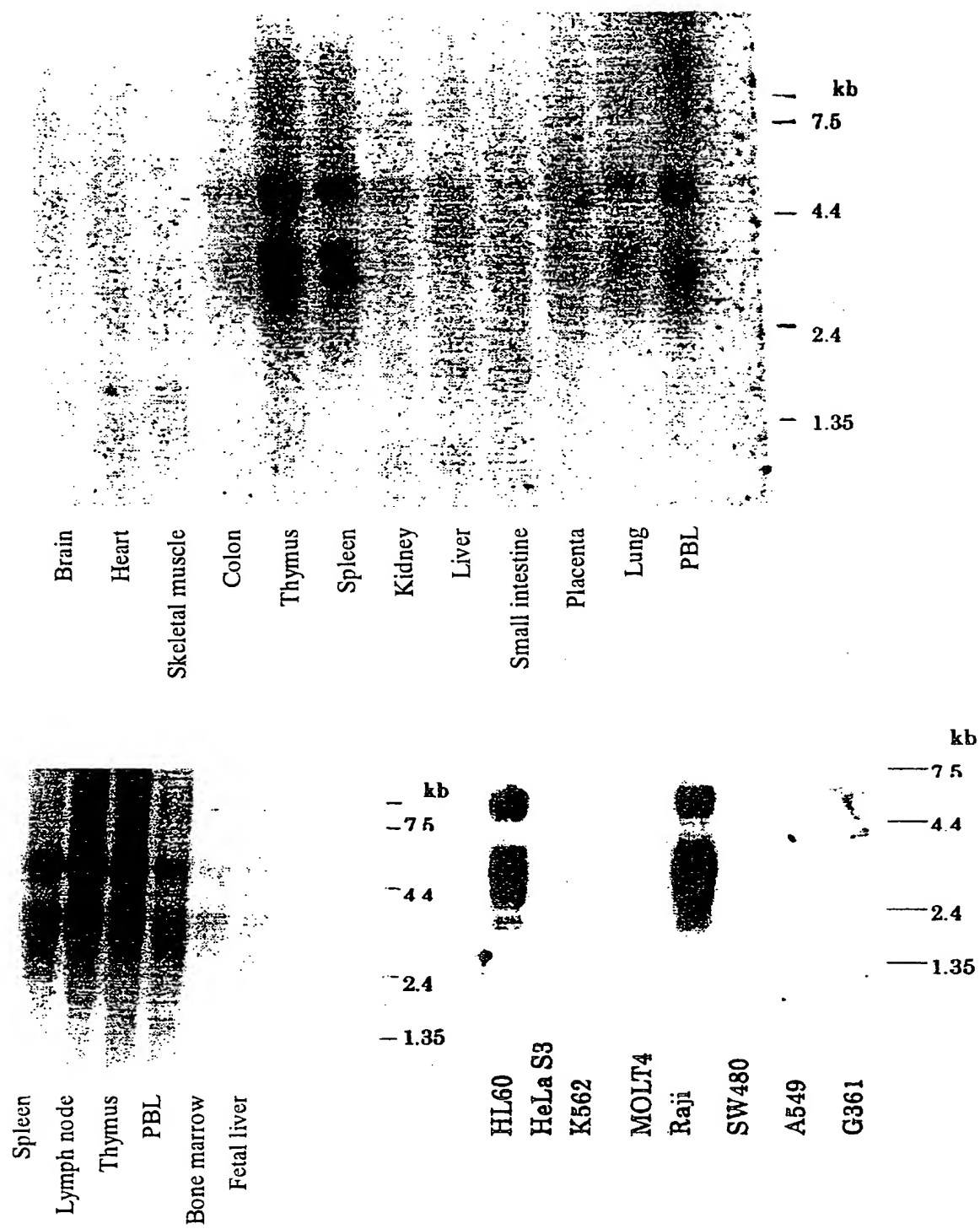
1050 1060 1070 1080 1090 1100 1110 1120
L I S V D S R S V S L L P L E F R K D S S Y E L Q V
CGGGCAGGGCCCATGCCTGGCTCCTCCTACCAGGGGACCTGGAGTGAATGGAGTGACCGGTATCTTTTCAGACCCAGTC

R A G P M P G S S Y Q G T W S E W S D P V I F Q T Q S

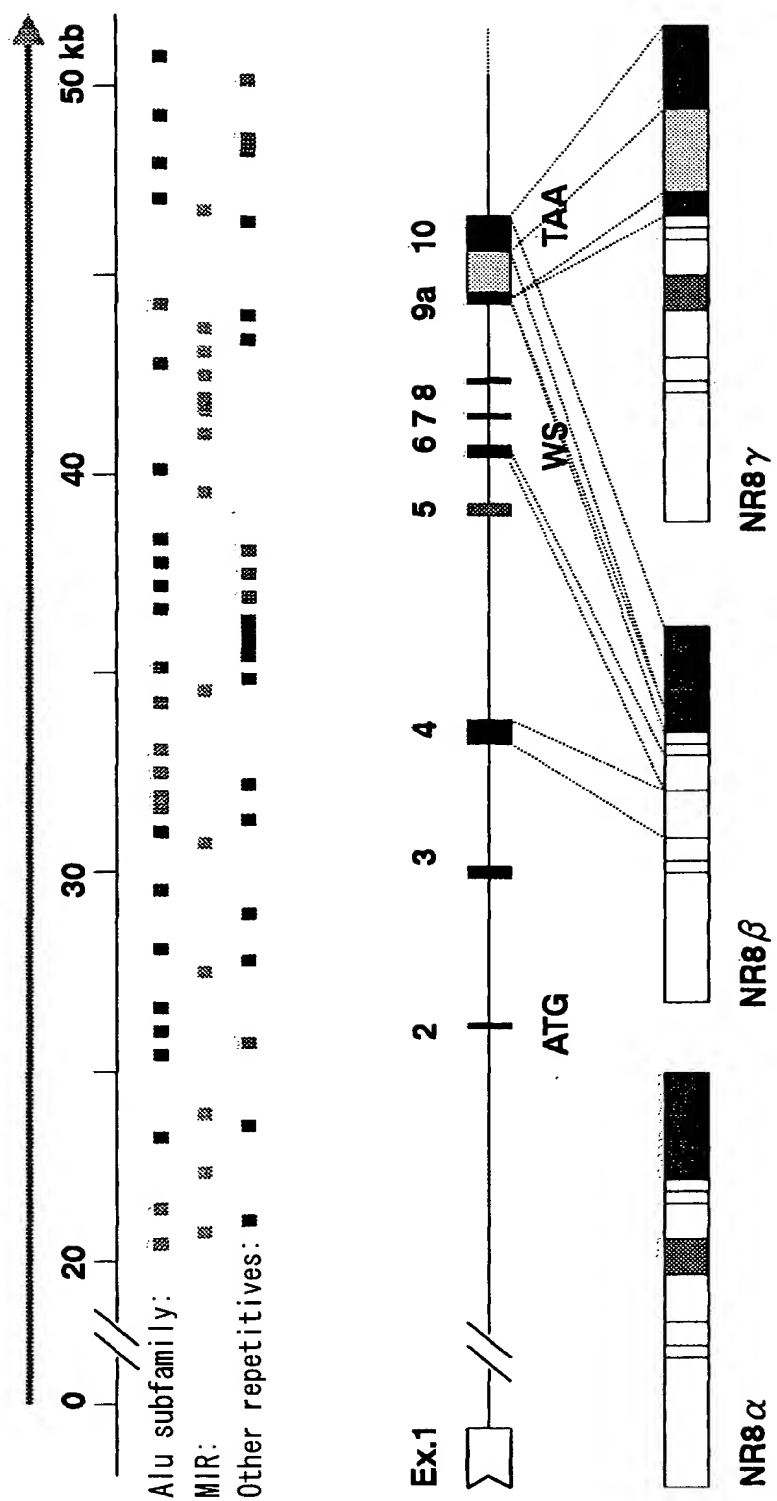
[Fig. 10]

1130 1140 1150 1160 1170 1180 1190 1200
 AGAGGAGTTAAAGGAAGGCTGGAACCCTCACCTGCTGCTTCTCCTCCTGCTTGTCATAGTCTTCATTCTGCTTCTGGA
 E E L K E G W N P H L L L L L L V I V F I P A F W S
 1210 1220 1230 1240 1250 1260 1270 1280
 GCCTGAAGACCCATCCATTGTGGAGGCTATGGAAGAAGATATGGGCCGTCCCAGCCCTGAGCGGTTCTTCATGCCCTG
 L K T H P L W R L W K K I W A V P S P E R F F M P L
 1290 1300 1310 1320 1330 1340 1350 1360
 TACAAGGGCTGCAGCGGAGACTTCAAGAAATGGGTGGGTGCACCCCTTCACTGGCTCCAGCCTGGAGCTGGGACCCTGGAG
 Y K G C S G D F K K W V G A P F T G S S L E L G P W S
 1370 1380 1390 1400 1410 1420 1430 1440
 CCCAGAGGTGCCCTCCACCCTGGAGGTGTACAGCTGCCACCCACCACGGAGCCCGCCAAGAGGCTGCAGCTCACGGAGC
 P E V P S T L E V Y S C H P P R S P A K R L Q L T E L
 1450 1460 1470 1480 1490 1500 1510 1520
 TACAAGAACCAGCAGAGCTGGTGGAGTCTGACGGTGTGCCAAGCCCAGCTTCTGGCCGACAGCCAGAAGCTCGGGGGGC
Q E P A E L V E S D G V P K P S F W P T A Q N S G G
 1530 1540 1550 1560 1570 1580 1590 1600
 TCAGCTTACAGTGAGGAGAGGGATCGGCCATACGGCCTGGTGTCCATTGACACAGTGACTGTGCTAGATGCAGAGGGGCC
S A Y S E E R D R P Y G L V S I D T V T V L D A E G P
 1610 1620 1630 1640 1650 1660 1670 1680
 ATGCACCTGGCCCTGCAGCTGTGAGGATGACGGCTACCCAGCCCTGGACCTGGATGCTGGCCTGGAGCCCAGCCAGGCC
C T W P C S C E D D G Y P A L D L D A G L E P S P G L
 1690 1700 1710 1720 1730 1740 1750 1760
 TAGAGGACCCACTCTTGATGCAGGGACCACAGTCCTGTCTGCTGGCTGTGTCTCAGCTGGCAGCCCTGGGCTAGGAGGG
E D P L L D A G T T V L S C G C V S A G S P G L G G
 1770 1780 1790 1800 1810 1820 1830 1840
 CCCCTGGGAAGCCTCCTGGACAGACTAAAGCCACCCCTTGCGATGGGGAGGACTGGGCTGGGGGACTGCCCTGGGGTGG
P L G S L L D R L K P P L A D G E D W A G G L P W G G
 1850 1860 1870 1880 1890 1900 1910 1920
 CCGGTCACCTGGAGGGGTCTCAGAGAGTGAGGCGGGCTCACCCCTGGCCGGCCTGGATATGGACACGTTTGACAGTGGCT
R S P G G V S E S E A G S P L A G L D M D T F D S G F
 1930 1940 1950 1960 1970 1980 1990 2000
 TTGTGGGCTCTGACTGCAGCAGCCCTGTGGAGTGTGACTTCACCAGCCCCGGGGACGAAGGACCCCCCGGAGCTACCTC
V G S D C S S P V E C D F T S P G D E G P P R S Y L
 2010 2020 2030 2040 2050 2060 2070 2080
 CGCCAGTGGGTGGTCATTCTCGCCACTTTCGAGCCCTGGACCCAGCCAGCTAATGAGGCTGACTGGATGTCCAGAG
 R Q W V V I P P P L S S P G P Q A S * *

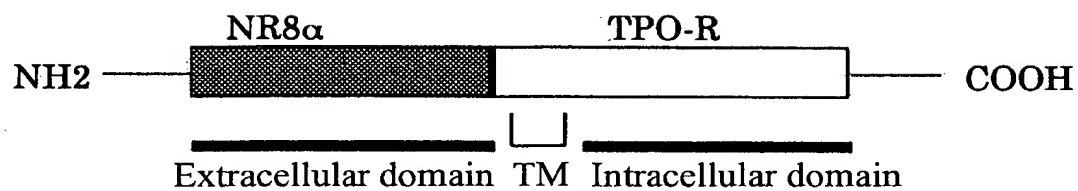
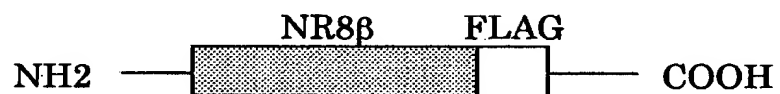
[Fig. 12]



AC002303



[Fig. 14]

pEF-NR8/ TPO-R**pEF-NR8/ IgG-Fc****pEF-BOS/ NR8b FLAG**

[Document Name] Abstract

[Abstract]

[Problems to Be Solved] The objective of the present invention is providing novel proteins, and genes encoding the proteins, and uses thereof.

5 [Means to Solve the Problems] The present invention provides novel hemopoietin receptor proteins (proteins comprising the amino acid sequence of SEQ ID NOs: 1, 3, 5, 7), proteins comprising a modified amino acid sequence of the amino acid sequence of the above protein
10 in which one or more amino acids have been deleted, added, and/or replaced with another amino acid, genes encoding these proteins, methods of producing the proteins, as well as uses of these proteins and genes.

[Selected Drawings] None

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☒ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.